COMPOSITIO MATHEMATICA

Level-raising for Saito–Kurokawa forms

Claus M. Sorensen

Compositio Math. **145** (2009), 915–953.

doi:10.1112/S0010437X09004084

FOUNDATION
COMPOSITIO
MATHEMATICA

The London
Mathematical
Society

Compositio Math. 145 (2009) 915–953
doi:10.1112/S0010437X09004084

Level-raising for Saito–Kurokawa forms

Claus M. Sorensen

Abstract

This paper provides congruences between unstable and stable automorphic forms for the symplectic similitude group GSp(4). More precisely, we raise the level of certain CAP representations Π arising from classical modular forms. We first transfer Π to π on a suitable inner form G; this is achieved by θ-lifting. For π, we prove a precise level-raising result that is inspired by the work of Bellaiche and Clozel and which relies on computations of Schmidt. We thus obtain a $\tilde{\pi}$ congruent to π, with a local component that is irreducibly induced from an unramified twist of the Steinberg representation of the Klingen parabolic. To transfer $\tilde{\pi}$ back to GSp(4), we use Arthur's stable trace formula. Since $\tilde{\pi}$ has a local component of the above type, all endoscopic error terms vanish. Indeed, by results due to Weissauer, we only need to show that such a component does not participate in the θ-correspondence with any GO(4); this is an exercise in using Kudla's filtration of the Jacquet modules of the Weil representation. We therefore obtain a cuspidal automorphic representation $\tilde{\Pi}$ of GSp(4), congruent to Π, which is neither CAP nor endoscopic. It is crucial for our application that we can arrange for $\tilde{\Pi}$ to have vectors fixed by the non-special maximal compact subgroups at all primes dividing N. Since G is necessarily ramified at some prime r, we have to show a non-special analogue of the fundamental lemma at r. Finally, we give an application of our main result to the Bloch–Kato conjecture, assuming a conjecture of Skinner and Urban on the rank of the monodromy operators at the primes dividing N.

1. Introduction

We fix a prime r and let D be the quaternion algebra over $\mathbb{Q}$ with ramification locus $S = \{\infty, r\}$. Let G be the unitary similitude group of D^2, where we take the hermitian form to be the identity matrix I. Thus, for example,

$$G(\mathbb{Q}) = \{x \in \mathrm{GL}(2, D) : x^* x = c(x) I,\ c(x) \in \mathbb{Q}^*\}.$$

Then, G is an inner form of GSp(4) such that $G(\mathbb{R})$ is compact modulo its center. More precisely, its adjoint group $G^{\mathrm{ad}}(\mathbb{R})$ is anisotropic SO(5). Similarly, $G^{\mathrm{ad}}(\mathbb{Q}_r)$ is the special orthogonal group of a quadratic form in five variables over $\mathbb{Q}_r$ with Witt index 1. There is another description of $G(\mathbb{Q}_r)$ in § 5.3. At all other primes p, the group G is split and hence $G(\mathbb{Q}_p)$ can be identified with $\mathrm{GSp}(4, \mathbb{Q}_p)$. Let $\mathbb{A}_f$ denote the finite part of the adeles $\mathbb{A}$.

The compact open subgroups K in $G(\mathbb{A}_f)$ form a directed set by opposite inclusion. Let $\mathcal{H}_{K,\mathbb{Z}}$ denote the natural $\mathbb{Z}$-structure in the Hecke algebra of K-biinvariant compactly supported functions on $G(\mathbb{A}_f)$. As K varies, the centers $Z(\mathcal{H}_{K,\mathbb{Z}})$ form an inverse system of algebras with respect to the canonical maps $Z(\mathcal{H}_{J,\mathbb{Z}}) \to Z(\mathcal{H}_{K,\mathbb{Z}})$ given by $\phi \mapsto e_K \star \phi$ for $J \subset K$. Consider the

Received 4 February 2008, accepted in final form 22 October 2008, published online 18 June 2009.
2000 Mathematics Subject Classification 11E57, 11E72, 11F27, 11F33, 11F41, 11F70, 11F72.
Keywords: automorphic forms, level-raising congruences, trace formula, Selmer groups.

inverse limit

$$\mathcal{Z} = \varprojlim Z(\mathcal{H}_{K,\mathbb{Z}}).$$

This makes sense locally, and $\mathcal{Z} = \otimes_{p<\infty} \mathcal{Z}_p$, where $\mathcal{Z}_p$ is obtained by the analogous construction at p. If π is an irreducible admissible representation of $G(\mathbb{A})$, there is a unique character $\eta_\pi : \mathcal{Z} \to \mathbb{C}$ such that $\eta_\pi = \eta_{\pi_f^K} \circ \mathrm{pr}_K$ whenever $\pi_f^K \neq 0$. Similarly, we have characters η_{π_p} locally, and then $\eta_\pi = \otimes_{p<\infty} \eta_{\pi_p}$ under the isomorphism above. If π is automorphic and $\pi_\infty = \mathbf{1}$, then the values of η_π are algebraic integers and we will use the following definition of being congruent.

Definition 1.1. Let $\tilde{\pi}$ and π be automorphic representations of $G(\mathbb{A})$, both trivial at infinity, and let λ be a finite place of $\bar{\mathbb{Q}}$. Then we define $\tilde{\pi}$ and π to be congruent modulo λ when the congruence

$$\eta_{\tilde{\pi}}(\phi) \equiv \eta_\pi(\phi) \pmod{\lambda}$$

holds for all $\phi \in \mathcal{Z}$. In this case, we will use the notation: $\tilde{\pi} \equiv \pi \pmod{\lambda}$.

Analogously, it makes sense to say that the local components $\tilde{\pi}_p$ and π_p are congruent. Then $\tilde{\pi} \equiv \pi \pmod{\lambda}$ if and only if $\tilde{\pi}_p \equiv \pi_p \pmod{\lambda}$ for all $p < \infty$. If $\tilde{\pi}_p$ and π_p are both unramified, $\tilde{\pi}_p \equiv \pi_p \pmod{\lambda}$ means that their (normalized) Satake parameters are congruent, up to permutation.

The following definition gives the analogue of the notion, from Clozel's paper [Clo00], of an automorphic representation being Eisenstein modulo λ.

Definition 1.2. Let π be an automorphic representation of $G(\mathbb{A})$ with $\pi_\infty = \mathbf{1}$, and let λ be a finite place of $\bar{\mathbb{Q}}$. We say that π is abelian modulo λ if there exists an automorphic character χ of $G(\mathbb{A})$ with infinity type $\chi_\infty = \mathbf{1}$ such that

$$\pi \equiv \chi \pmod{\lambda}.$$

We prefer the terminology 'abelian modulo λ' since the group G has no $\mathbb{Q}$-parabolics and hence no Eisenstein series. We note that there *exists* non-abelian π exactly because $G(\mathbb{R})$ is assumed to be compact modulo its center.

For the next theorem, we fix a good small compact open subgroup K of the form $\prod K_p$ (see § 2.1.4 for the precise definition of 'good small'). Let N be an integer such that $p \nmid N$ implies that K_p is hyperspecial. Then we have the following main result on level-raising for the inner form G.

Theorem A. *Let $\lambda | \ell$ be a finite place of $\bar{\mathbb{Q}}$, with ℓ outside a finite set determined by K. Let π be an automorphic representation of $G(\mathbb{A})$, with ω_π and π_∞ trivial, such that $\pi_f^K \neq 0$. Assume that π occurs with multiplicity one and that π is non-abelian modulo λ. Suppose $q \nmid N\ell$ is a prime number with $q^i \neq 1 \pmod{\ell}$ for $i = 1, \ldots, 4$ such that, modulo the Weyl-action, we have the congruence*

$$\mathfrak{t}_{\pi_q \otimes |c|^{-3/2}} \equiv \begin{pmatrix} 1 & & & \\ & q & & \\ & & q^2 & \\ & & & q^3 \end{pmatrix} \pmod{\lambda},$$

where $\mathfrak{t}$ *denotes the Satake parameter of the lift to* $\mathrm{GL}(4)$*. Then there exists an automorphic representation* $\tilde{\pi} \equiv \pi \pmod{\lambda}$ *of* $G(\mathbb{A})$*, with* $\omega_{\tilde{\pi}}$ *and* $\tilde{\pi}_\infty$ *trivial, such that* $\tilde{\pi}_f^{K^q} \neq 0$ *and*

$$\tilde{\pi}_q \text{ is of type } \begin{cases} \text{IIIa} \\ \text{IIa} \end{cases} \quad \text{when } \pi_q \text{ is of type } \begin{cases} \text{IIb} \\ \text{IIIb}, \end{cases}$$

respectively. In the remaining case where π_q *is generic, one can choose a* $\tilde{\pi}_q$ *of type IIa and a* $\tilde{\pi}_q$ *of type IIIa (one for each type).*

The finite set of primes ℓ that we have to discard are those dividing the discriminant of the Hecke algebra of K; see § 2.1.4 below for more details.

We use the classification of [Sch05], which is reproduced in Appendix A and Appendix B. Note that the two types IIb and IIIb are the typical unramified local components of CAP representations. The representations of type IIa are of the form $\chi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$ (induced from the Siegel parabolic), while those of type IIIa are $\chi \rtimes \sigma\mathrm{St}_{\mathrm{GL}(2)}$ (induced from the Klingen parabolic). Both are ramified, generic, and Klingen- and Siegel-spherical. Moreover, a representation of type IIIa is tempered if and only if it is unitary. In the generic case, one can choose between the types IIa and IIIa depending on the application one has in mind. Representations of type IIa are expected to transfer to the inner form over $\mathbb{Q}_q$, while those of type IIIa cannot occur in endoscopic lifts. We will prove this below.

The proof of the above theorem is inspired by the work of Bellaiche [Bel02] and Clozel [Clo00]. They were both dealing with a unitary group $\mathrm{U}(3)$, split over some imaginary quadratic extension $E/\mathbb{Q}$. Clozel considered the case where q is inert in E; here the semisimple rank is one, and he obtained a $\tilde{\pi}$ with a Steinberg component at q. In his thesis, Bellaiche dealt with the case where q is split in E; here the semisimple rank is two, which makes things more complicated. In this case, one gets a $\tilde{\pi}$ with $\tilde{\pi}_q$ ramified but having fixed vectors under any maximal parahoric in $\mathrm{GL}(3)$. This, in turn, implies that $\tilde{\pi}_q = \chi\mathrm{St}_{\mathrm{GL}(2)} \times \sigma$ by the classification of Iwahori-spherical representations of $\mathrm{GL}(3)$ in [Sor06]. For $\mathrm{GSp}(4)$ this classification is much more complicated, but fortunately it has been done by Schmidt [Sch05]. To really utilize the tables in [Sch05], we need to modify Bellaiche's argument a bit. For example, we incorporate the action of the Bernstein center at q, and we get a precise condition on what characteristics ℓ we need to discard. At a crucial point, we rely on results of Lazarus [Laz00] describing the structure of universal modules.

The approach in [Sor06] is different. There, we use the arguments of Taylor [Tay89] in a more general setup. In the special case of an inner form of $\mathrm{GSp}(4)$, the main result of [Sor06] has weaker assumptions (no multiplicity-one or banality condition is needed, and it works for arbitrary ℓ) but also a weaker conclusion (one can only say that $\tilde{\pi}_q$ is of type I, IIa *or* IIIa). In particular, $\tilde{\pi}_q$ could be a full unramified principal series, and we would be unable to rule out the possibility that $\tilde{\pi}$ is endoscopic. Therefore, the results of [Sor06] are too weak for the applications to the Bloch–Kato conjecture which we have in mind.

Next, we prove a purely local result at the prime r, which will be crucial later on for our application of the trace formula. Now D_r is the division quaternion $\mathbb{Q}_r$-algebra, and we fix an unramified quadratic subfield E. Let θ be the generator for $\mathrm{Gal}(E/\mathbb{Q}_r)$. From now on, let G' denote $\mathrm{GSp}(4)$. We view G as the non-split inner form of G' over $\mathbb{Q}_r$. It comes with a class of inner twistings

$$\psi : G \to G',$$

that is, ψ is an isomorphism over $\bar{\mathbb{Q}}_r$ such that $\sigma\psi \circ \psi^{-1}$ is an inner automorphism of G' for all σ in the Galois group of $\mathbb{Q}_r$. We fix a ψ defined over E. Since G^{der} is simply connected, stable conjugacy is just $G(\bar{\mathbb{Q}}_r)$-conjugacy; similarly for G'. Then ψ defines an injection from the semisimple stable conjugacy classes in $G(\mathbb{Q}_r)$ to the semisimple stable conjugacy classes in $G'(\mathbb{Q}_r)$. Now, we introduce orbital integrals: let $\gamma \in G(\mathbb{Q}_r)$ be a semisimple element, and let

$$G_\gamma(\mathbb{Q}_r) = \{x \in G(\mathbb{Q}_r) : x^{-1}\gamma x = \gamma\}.$$

Choose Haar measures on $G(\mathbb{Q}_r)$ and $G_\gamma(\mathbb{Q}_r)$, and consider the orbital integral

$$O_\gamma(f) = \int_{G_\gamma(\mathbb{Q}_r)\backslash G(\mathbb{Q}_r)} f(x^{-1}\gamma x)\, dx$$

of a function $f \in C_c^\infty(G^{\mathrm{ad}}(\mathbb{Q}_r))$. Take $\{\tilde{\gamma}\}$ to be a set of representatives for the conjugacy classes within the stable conjugacy class of γ, modulo the center. Then $G_{\tilde{\gamma}}$ is an inner form of G_γ, and we choose compatible measures. Let $e(G_{\tilde{\gamma}})$ denote the Kottwitz sign [Kot83] and form the *stable* orbital integral

$$SO_\gamma(f) = \sum_{\tilde{\gamma}} e(G_{\tilde{\gamma}}) O_{\tilde{\gamma}}(f).$$

The definitions for G' are, of course, completely analogous. Now consider two functions $f \in C_c^\infty(G^{\mathrm{ad}}(\mathbb{Q}_r))$ and $f' \in C_c^\infty(G'^{\mathrm{ad}}(\mathbb{Q}_r))$. They are said to have matching orbital integrals if and only if for all semisimple elements $\gamma' \in G'(\mathbb{Q}_r)$,

$$SO_{\gamma'}(f') = \begin{cases} SO_\gamma(f) & \text{if } \gamma' \text{ belongs to } \psi(\gamma), \\ 0 & \text{if } \gamma' \text{ does not come from } G. \end{cases}$$

Here we use compatible Haar measures on both sides. We note that Waldspurger has shown in [Wal97] (using results of Langlands and Shelstad [LS90]) that one can always find a function f' matching a given f. We take f to be the idempotent of a maximal compact subgroup in $G(\mathbb{Q}_r)$ and show that we can take f' to be biinvariant under a corresponding maximal compact subgroup in $G'(\mathbb{Q}_r)$.

The semisimple $\mathbb{Q}_r$-rank of G' is two, and the (reduced) building is covered by two-dimensional apartments. Each apartment is tessellated by equilateral right-angled triangles. The vertices at the right angles are non-special, whereas all the other vertices are hyperspecial. Correspondingly, the group $G'(\mathbb{Q}_r)$ has two conjugacy classes of maximal compact subgroups: the hyperspecial subgroups and the so-called (non-special) paramodular subgroups.

The group G has semisimple $\mathbb{Q}_r$-rank one, so its (reduced) building is an inhomogeneous tree. In fact, for $r = 2$ there is a picture of it in the article of Tits [Tit79, p. 48]. All of its vertices are special. Each edge has one vertex of order $r^2 + 1$ and one vertex of order $r + 1$. The former maps to a non-special vertex in the building over E, whereas the latter maps to the midpoint of a long edge. The stabilizer of a vertex of order $r^2 + 1$ is also called paramodular.

Let K' be a paramodular group in $G'(\mathbb{Q}_r)$. Concretely, one can take K' to be the subgroup generated by the Klingen parahoric and the matrix

$$\begin{pmatrix} & & & -r^{-1} \\ & 1 & & \\ & & 1 & \\ r & & & \end{pmatrix}.$$

Besides Z and K' itself, its normalizer contains an element η, called the Atkin–Lehner element in the paper of Schmidt [Sch05]. It has the following form:

$$\eta = \begin{pmatrix} & & 1 & \\ & & & 1 \\ r & & & \\ & r & & \end{pmatrix}.$$

Note that it satisfies the identity $\eta^2 = rI$. Our main matching result is then given in the following theorem.

Theorem B. *Let K and K' be arbitrary paramodular subgroups in $G(\mathbb{Q}_r)$ and $G'(\mathbb{Q}_r)$, respectively. Then the characteristic functions e_K and $e_{\eta K'}$ have matching orbital integrals in the sense described above.*

Here e_K, for example, denotes the characteristic function of ZK. The main ingredient of the proof is a slight modification of the results obtained by Kottwitz in [Kot86]. First, since G splits over E, we may compare the stable orbital integrals of e_K and $e_{\eta K'}$ to stable *twisted* orbital integrals on $G(E)$ and $G'(E)$. In turn, these integrals can be compared explicitly by hand using the inner twisting ψ.

The next result is a special case of the Langlands functoriality conjecture. More precisely, it is an analogue for GSp(4) of the Jacquet–Langlands correspondence between the spectra of GL(2) and its inner forms. It allows us to transfer the $\tilde{\pi}$ from Theorem A to $G'(\mathbb{A})$ in the cases we are interested in. The notion of being endoscopic is made precise in § 4.1.4. Here we note that a cuspidal automorphic representation Π of $G'(\mathbb{A})$ is said to be CAP with respect to a $\mathbb{Q}$-parabolic P, with Levi component M, if there exists a cuspidal automorphic representation τ of $M(\mathbb{A})$ such that Π is weakly equivalent to the constituents of the induced representation of τ to $G'(\mathbb{A})$. Recall that weakly equivalent means isomorphic at all but finitely many places.

Theorem C. *Let π be an automorphic representation of $G(\mathbb{A})$, with ω_π and π_∞ trivial. Suppose there exists a prime $q \notin S$ such that π_q is of type IIIa of the form $\chi \rtimes \sigma St_{\mathrm{GL}(2)}$ where $\chi^2 \neq \mathbf{1}$. Pick a cohomological discrete series representation $\Pi_{\mathbf{1}}$ of $G'(\mathbb{R})$, holomorphic or generic. Then there exists a cuspidal automorphic representation Π of $G'(\mathbb{A})$, with ω_Π trivial and $\Pi_\infty = \Pi_{\mathbf{1}}$, such that $\Pi_p = \pi_p$ for all $p \notin S$. Any such Π is neither CAP nor endoscopic. Moreover, if π_r is para-spherical (that is, has vectors fixed by a paramodular group), then there exists a Π as above with Π_r para-spherical and ramified.*

Let us briefly sketch the ideas of the proof. The main tool is Arthur's stable trace formula. The point is that the endoscopic group $\mathrm{PGL}(2) \times \mathrm{PGL}(2)$ for PGSp(4) has no endoscopy itself, and therefore we only need the standard fundamental lemma proved by Hales (not the weighted version). Hales has also computed the Shalika germs for GSp(4) and its inner forms. Then, from the general results of Langlands and Shelstad on descent for transfer factors, one immediately deduces the transfer conjecture in our case. In fact, more recently, Waldspurger has shown in general that the fundamental lemma implies the transfer conjecture. Intuitively, this enables us to match the geometric sides of the trace formulas for G and G'. Consequently, the spectral sides match and we can compare the spectra. However, there is a serious problem to overcome: the distribution defined by the trace formula is unstable. One makes it stable by subtracting suitable endoscopic error terms. To show that these error terms vanish in our situation, we invoke results of Weissauer describing endoscopic lifts in terms of the θ-correspondence with $\mathrm{GO}(X)$ for four-dimensional X. It remains to show that type-IIIa representations do not participate in these correspondences. For this purpose, we use Kudla's filtration of the Jacquet modules of the Weil representation. Roughly, this filtration reveals that the Weil representation is compatible with parabolic induction. Thus, we are reduced to showing that the Steinberg representation $\mathrm{St}_{\mathrm{SL}(2)}$ does not occur in the θ-correspondence with split $\mathrm{O}(2)$; this is a well-known fact. A standard argument, based on the linear independence of characters, then gives a discrete automorphic representation Π with $\Pi^S = \pi^S$. It is actually cuspidal: by the theory of Eisenstein series we can rule out that it occurs in the residual spectrum, since it has a tempered component. To see that it is not CAP, we make use of work due to Piatetski-Shapiro and Soudry. To make sure the component Π_∞ lies in the expected L-packet and that we can indeed choose a specific member, we rely on the exhaustive work of Shelstad in the archimedean case.

To get the paramodular refinement at r, we appeal to Theorem B. In fact, we get a Π such that the Atkin–Lehner operator on the paramodular invariants of Π_r has a positive trace. Using work of Weissauer [Wei05] on the Ramanujan conjecture, we show that Π_r is in fact also ramified. Then, by the computations of Schmidt, Π_r must be of type IIa, Vb, Vc or VIc. We expect that Π_r is necessarily tempered. If so, it is of type IIa of the form $\chi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$ with $\chi\sigma$ non-trivial quadratic.

Seemingly, slight modifications of the above trace-formula argument, combined with Weissauer's work on weak endoscopic lifts, should allow one to prove the existence of weak Jacquet–Langlands transfer in general for $\mathrm{GSp}(4)$.

The foregoing discussion culminates in the following main result, which provides congruences between unstable and stable automorphic forms for $\mathrm{GSp}(4)$. Let $f \in S_4(\Gamma_0(N))$ be a newform of weight 4 and square-free level N (and trivial character). The assumption on the weight is made to simplify things as much as possible. Indeed, the Saito–Kurokawa lift (which we introduce below) will be cohomological, so that its transfer to an inner form will be trivial at infinity. Hence, the algebraic modular forms we end up looking at are just functions on a finite set. We assume throughout that f has root number

$$\epsilon_f = -1;$$

in other words, the L-function $L(s, f)$ vanishes to an odd order at $s = 2$. For example, such a newform f exists for $N = 13$. By this sign condition, we may lift f to a Saito–Kurokawa form $\mathrm{SK}(f)$ on $\mathrm{GSp}(4)$. This is a CAP representation, holomorphic at infinity and having Galois representation

$$\rho_{\mathrm{SK}(f),\lambda} \simeq \rho_{f,\lambda} \oplus \omega_\ell^{-1} \oplus \omega_\ell^{-2},$$

where ω_ℓ is the ℓ-adic cyclotomic character and $\rho_{f,\lambda}$ is the system of Galois representations attached to f by Deligne [Del71]. More recently, Laumon [Lau97] and Weissauer [Wei05] have attached Galois representations to *any* cuspidal automorphic representation of $\mathrm{GSp}(4)$, which is a discrete series at infinity. We produce congruences between $\mathrm{SK}(f)$ and certain stable forms of small level.

Theorem D. *With notation as above, let $\lambda|\ell$ be a finite place of $\bar{\mathbb{Q}}$, with ℓ outside a finite set of primes determined by N, such that $\bar{\rho}_{f,\lambda}$ is irreducible. Suppose $q \nmid N\ell$ is a prime such that:*

- $q^i \not\equiv 1 \pmod{\ell}$ *for* $i = 1, \ldots, 4$;
- $\bar{\rho}_{f,\lambda}(\mathrm{Frob}_q)$ *has a fixed vector.*

Then there exists a cuspidal automorphic representation $\Pi \equiv \mathrm{SK}(f) \pmod{\lambda}$ *of* $\mathrm{PGSp}(4)$ *which is neither* CAP *nor endoscopic and such that* Π_∞ *is the cohomological holomorphic discrete series representation,* Π_p *is unramified and tempered for* $p \nmid Nq$*, and:*

- *the Galois representation* $\rho_{\Pi,\lambda}$ *is irreducible;*
- Π_q *is of type IIIa (hence tempered, generic and ramified);*
- Π_p *is para-spherical for all primes* p *dividing* N*.*

Moreover, if f *is not CM, there exists a positive density of such primes* q*.*

The proof is a combination of all of our previous results. Let us outline the main ideas. Since f has trivial character, it generates a cuspidal automorphic representation τ of $\mathrm{GL}(2,\mathbb{A})$ with trivial central character and with τ_∞ being the holomorphic discrete series of weight 4, so that we have the equality $L(s-3/2,\tau_f) = L(s,f)$ between L-functions. We may, and will, view τ as a cuspidal automorphic representation of $\mathrm{PGL}(2)$. Choose a prime r such that τ_r is the Steinberg representation (and *not* its unramified quadratic twist). Let G be the definite inner form of $\mathrm{GSp}(4)$ with ramification locus $\{\infty, r\}$; its adjoint group G^{ad} is the special orthogonal group of a definite quadratic form in five variables over $\mathbb{Q}$. Let A_τ be the global Waldspurger packet for the metaplectic group $\widetilde{\mathrm{SL}}(2)$ determined by τ. Then $\mathrm{SK}(f) = \theta(\sigma)$ for some $\sigma \in A_\tau$. We consider the reflection $\breve{\sigma} \in A_\tau$ and its lifting $\theta(\breve{\sigma})$ to the inner form G. This turns out to be para-spherical at all primes dividing N. By Theorem A, we can raise the level: since $\mathrm{SK}(f)$ has local components of type IIb outside N, we get a $\pi \equiv \theta(\breve{\sigma}) \pmod{\lambda}$ with π_q of type IIIa. Then, by Theorem C, we can transfer π to an automorphic representation Π of $\mathrm{GSp}(4)$ agreeing with π outside of $\{\infty, r\}$. The irreducibility of $\rho_{\Pi,\lambda}$ was essentially proved by Ramakrishnan in [Ram08].

Finally, we give an application of Theorem D to proving new cases of the Bloch–Kato conjecture for classical modular forms, upon assuming a conjecture of Skinner and Urban. Before stating our result, we briefly recall the definition of Selmer groups. Let V be a continuous representation of the Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, with coefficients in a finite extension $L/\mathbb{Q}_\ell$. Choose a lattice Λ and define W by

$$0 \to \Lambda \xrightarrow{i} V \xrightarrow{\mathrm{pr}} W \to 0.$$

Let λ be the maximal ideal in the ring of integers of L. Then we identify the reduction $\Lambda/\lambda\Lambda$ with the λ-torsion in W. For each prime p, let I_p be the inertia group $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p^{\mathrm{nr}})$. Let B_{cris} be Fontaine's crystalline Barsotti–Tate ring [BK90]. Then we define the *finite* part of the local Galois cohomology to be

$$H^1_f(\mathbb{Q}_p, V) = \begin{cases} \ker\{H^1(\mathbb{Q}_p, V) \to H^1(\mathbb{Q}_p^{\mathrm{nr}}, V)\} & \text{for } p \neq \ell, \\ \ker\{H^1(\mathbb{Q}_\ell, V) \to H^1(\mathbb{Q}_\ell, B_{\mathrm{cris}} \otimes V)\} & \text{for } p = \ell. \end{cases}$$

The Selmer group $H^1_f(\mathbb{Q}, V)$ is then cut out by these local conditions:

$$H^1_f(\mathbb{Q}, V) = \ker\left\{ H^1(\mathbb{Q}, V) \to \prod_p H^1(\mathbb{Q}_p, V)/H^1_f(\mathbb{Q}_p, V) \right\}.$$

Using the maps induced by i and pr, we define the finite parts for Λ and W:

$$H^1_f(\mathbb{Q}_p, \Lambda) = i_*^{-1} H^1_f(\mathbb{Q}_p, V), \quad H^1_f(\mathbb{Q}_p, W) = \mathrm{pr}_* H^1_f(\mathbb{Q}_p, V).$$

The Selmer groups $H^1_f(\mathbb{Q}, \Lambda)$ and $H^1_f(\mathbb{Q}, W)$ are then defined as above. If V is the ℓ-adic realization of a motive, the latter group is sometimes called the ℓ-part of the Selmer group of the motive. It sits in a short exact sequence, where the quotient is conjecturally a finite ℓ-group:

$$0 \to \mathrm{pr}_* H^1_f(\mathbb{Q}, V) \to H^1_f(\mathbb{Q}, W) \to \mathrm{III}(\mathbb{Q}, W) \to 0.$$

This quotient $\mathrm{III}(\mathbb{Q}, W)$ is called the ℓ-part of the Tate–Shafarevich group. For the moment, let $\bar{\Lambda}$ denote the reduction $\Lambda/\lambda\Lambda$. Then the finite part $H^1_f(\mathbb{Q}_p, \bar{\Lambda})$ is defined to be the image of $H^1_f(\mathbb{Q}_p, \Lambda)$ under the natural map [Rub00, p. 17]. The Selmer group $H^1_f(\mathbb{Q}, \bar{\Lambda})$ is then defined as above. In the situations we will be interested in, it can be identified with the λ-torsion in $H^1_f(\mathbb{Q}, W)$.

The classes in $H^1_f(\mathbb{Q}, V)$ correspond to equivalence classes of certain extensions of the trivial representation $\mathbf{1}$ by V having *good* reduction. To make this precise, consider an extension of ℓ-adic $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$-modules

$$0 \to V \to X \to \mathbf{1} \to 0.$$

It is said to have good reduction at $p \neq \ell$ if the sequence remains exact after taking I_p-invariants. In particular, if V is unramified at p, this simply means that X is unramified at p. Similarly, the extension is said to have good reduction at ℓ if the sequence remains exact after applying the crystalline functor

$$D_{\mathrm{cris}}(V) = H^0(\mathbb{Q}_\ell, B_{\mathrm{cris}} \otimes V).$$

This is a filtered module of dimension at most $\dim_{\mathbb{Q}_\ell}(V)$. If the dimensions are equal, V is called crystalline. In this case, X is required to be crystalline. An extension X with good reduction everywhere gives rise to a cohomology class in $H^1_f(\mathbb{Q}, V)$ via the connecting homomorphism; this defines a bijection. The other Selmer groups above have similar interpretations. For example, the finite part $H^1_f(\mathbb{Q}_\ell, \bar{\Lambda})$ is related to the notion of being Fontaine–Laffaille [FL82].

Here we will be content with formulating the Bloch–Kato conjecture [BK90] for classical modular forms. At first, consider an arbitrary newform $f \in S_{2\kappa}(\Gamma_0(N))$. We take V above to be the κth Tate twist $\rho_{f,\lambda}(\kappa)$ of the Galois representation attached to f. Then, conjecturally, one has the following identity:

$$\mathrm{ord}_{s=\kappa} L(s, f) \stackrel{?}{=} \dim_{\mathbb{Q}_\ell} H^1_f(\mathbb{Q}, \rho_{f,\lambda}(\kappa)).$$

If $\epsilon_f = -1$, the L-function vanishes at the point $s = \kappa$ and so the conjecture predicts that the pertinent Selmer group is non-trivial. This was proved by Skinner and Urban in [SU06] under the assumption that f is ordinary at λ (meaning that the Hecke eigenvalue $a_\ell(f)$ is a λ-adic unit). Their proof relies on the deep results of Kato [Kat04]. However, in the square-free case they give a different argument, bypassing the work of Kato but instead relying on Conjecture 1 below.

Let ρ be a continuous representation of $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ on a finite-dimensional vector space V over the ℓ-adic field L. Assume $p \neq \ell$. Then, by a famous result of Grothendieck, ρ is potentially semistable. This means that there exists a nilpotent endomorphism $N : V \to V$ such that

$$\rho(\sigma) = \exp(t_\ell(\sigma) N)$$

for σ in a finite-index subgroup of I_p. Here $t_\ell : I_p \twoheadrightarrow \mathbb{Z}_\ell$ is a homomorphism intertwining the natural actions of the Weil group at p. The endomorphism N is called the monodromy operator. The following is basically [SU06, p. 41, Conjecture 3.1.7].

Conjecture 1. Let Π be a cuspidal automorphic representation of GSp(4) which is neither CAP nor endoscopic, with Π_∞ cohomological. Suppose the local component Π_p has non-zero vectors fixed by the paramodular group. Then the corresponding monodromy operator at p has rank at most one.

This conjecture follows from the expected compatibility between the local and global Langlands correspondences for GSp(4). Our application to the Bloch–Kato conjecture is contingent on Conjecture 1.

Theorem E. *Let $f \in S_4(\Gamma_0(N))$ be a newform of square-free level N with root number $\epsilon_f = -1$. Assume that f is not of CM type. Let $\lambda|\ell$ be a finite place of $\bar{\mathbb{Q}}$, with ℓ outside a finite set, such that $\bar{\rho}_{f,\lambda}$ is irreducible. Assume Conjecture 1 above. Then the mod-λ Selmer group $H^1_f(\mathbb{Q}, \bar{\rho}_{f,\lambda}(2))$ is non-trivial as predicted by the Bloch–Kato conjecture, since $L(s, f)$ vanishes at $s = 2$.*

By the result of Jordan and Livne on level-lowering for modular forms of higher weight [JL89], we may assume that $\bar{\rho}_{f,\lambda}$ is ramified at all primes $p|N$. Indeed, congruent eigenforms have equal root numbers (look at the W-operator). This minimality assumption turns out to be crucial for the proof.

It follows immediately from Theorem E that the Selmer group of $\rho_{f,\lambda}(2)$ is non-trivial, assuming the ℓ-part of the Tate–Shafarevich group is trivial. This should always be the case according to [BK90, p. 376, Conjecture 5.15].

We outline the main ideas of the proof of Theorem E. By Theorem D, we obtain a prime q and an automorphic representation Π. First, we choose a lattice Λ in the space of $\rho_{\Pi,\lambda}$ such that $\bar{\Lambda}$ has $\bar{\rho}_{f,\lambda}$ as its unique irreducible quotient. The goal is then to show that $\bar{\omega}_\ell^{-2}$ embeds into $\bar{\Lambda}$. If not, then $\bar{\omega}_\ell^{-1}$ would be the unique irreducible subrepresentation of $\bar{\Lambda}$. Thus we would get two non-split extensions

$$0 \to \bar{\omega}_\ell^{-1} \to X \to \bar{\omega}_\ell^{-2} \to 0 \quad \text{and} \quad 0 \to \bar{\omega}_\ell^{-2} \to Y \to \bar{\rho}_{f,\lambda} \to 0.$$

Both X and Y are subquotients of the étale intersection cohomology

$$IH^3_{\text{et}}(\bar{S}_K \times_{\mathbb{Q}} \bar{\mathbb{Q}}, \bar{\mathbb{Q}}_\ell),$$

where K is paramodular at primes dividing N, Klingen at q, and hyperspecial outside Nq. We denote by $\bar{S}_K$ the Satake compactification of the Siegel threefold S_K. Obviously, X and Y are then both unramified outside Nq. In addition, they are both Fontaine–Laffaille. To show that X and Y both have good reduction at primes dividing N, we use Conjecture 1 and our minimality assumption. At q the monodromy operator has order two by a result of Genestier and Tilouine; this allows us to show that X or Y is unramified at q, which is a contradiction. Indeed, by Kummer theory, $H^1_f(\mathbb{Q}, \bar{\omega}_\ell) = 0$; and, by Kato's paper [Kat04], we also have $H^1_f(\mathbb{Q}, \bar{\rho}_{f,\lambda}(1)) = 0$. Therefore $\bar{\omega}_\ell^{-2}$ does embed into $\bar{\Lambda}$, and we get a non-split extension with good reduction everywhere:

$$0 \to \bar{\omega}_\ell^{-1} \to Z \to \bar{\rho}_{f,\lambda} \to 0.$$

Twisting the dual extension, we obtain the desired class in $H^1_f(\mathbb{Q}, \bar{\rho}_{f,\lambda}(2))$.

Theorem E, and its proof, should be compared to the main theorem in Bellaiche's thesis [Bel02, Theorem VIII.1.7.2]. The latter gives the complete analogue for certain algebraic Hecke characters of an imaginary quadratic field by studying their endoscopic lifts to U(3), following Rogawski.

2. Level-raising

2.1 Algebraic modular forms

2.1.1 *The complex case.* Let G be an inner form of GSp(4) over $\mathbb{Q}$ such that $G(\mathbb{R})$ is compact modulo its center. Concretely, G is the unitary similitude group of D^2, where we take the hermitian form to be the identity I, and where D is some definite quaternion algebra over $\mathbb{Q}$ with ramification locus $S = \{\infty, r\}$ for some prime r. Let $c : G \to \mathbb{G}_m$ denote the similitude, and let $Z \simeq \mathbb{G}_m$ be the center. Then, let

$$\mathcal{A} = \{\text{smooth } f : Z(\mathbb{A}_f)G(\mathbb{Q})\backslash G(\mathbb{A}_f) \to \mathbb{C}\}.$$

Here $\mathbb{A}_f$ is the finite part of the adele ring $\mathbb{A}$. There is an admissible representation r of $G(\mathbb{A}_f)$ on this space given by right translations. In turn, the Hecke algebra $\mathcal{H}$ of compactly supported smooth functions on $G(\mathbb{A}_f)$ also acts in the usual way. We equip $\mathcal{A}$ with the pairing defined by the integral

$$\langle f, f' \rangle = \int_{G(\mathbb{Q})\backslash G(\mathbb{A}_f)} f(x) f'(x)\, dx.$$

This is well-defined since $G(\mathbb{Q})$ is a discrete cocompact subgroup of $G(\mathbb{A}_f)$, and we have

$$\langle r(\phi)f, f' \rangle = \langle f, r(\phi^\vee) f' \rangle,$$

where the anti-involution $\phi \mapsto \phi^\vee$ of $\mathcal{H}$ is defined by $\phi^\vee(x) = \phi(x^{-1})$; it reflects taking the contragredient. Now, let K be a compact open subgroup of $G(\mathbb{A}_f)$, and let $\mathcal{A}_K$ be the space of K-invariants. The Hecke algebra $\mathcal{H}_K$ of K-biinvariant compactly supported functions on $G(\mathbb{A}_f)$ then acts semisimply:

$$\mathcal{A}_K \simeq \bigoplus_{\Pi} m(\Pi)\Pi_f^K,$$

where Π varies over the automorphic representations of $G(\mathbb{A})$, with trivial central character, such that Π_∞ is trivial and $\Pi_f^K \neq 0$. Let us choose a complete set of representatives $\{x_i\}$ for the following finite set of cardinality h,

$$G(\mathbb{Q})\backslash G(\mathbb{A}_f)/K.$$

Then the map $f \mapsto (f(x_i))$ identifies $\mathcal{A}_K$ with a subspace of $\mathbb{C}^h$. Let us introduce

$$\Gamma_i = G(\mathbb{Q}) \cap x_i K x_i^{-1}.$$

These finite groups are all trivial if K is sufficiently small (to be precise, if the projection of K to some $G(\mathbb{Q}_p)$ does not contain non-trivial elements of finite order). Finally, let us consider the pairing on $\mathcal{A}$ restricted to $\mathcal{A}_K$. A straightforward calculation shows that we have the simple formula

$$\langle f, f' \rangle = \sum_i f(x_i) f'(x_i) \# \Gamma_i^{-1} \quad \forall f, f' \in \mathcal{A}_K,$$

up to normalization. In particular, it follows that the pairing is non-degenerate.

2.1.2 *Models over number fields.* Let $\mathcal{H}_{K,\mathbb{Z}}$ denote the natural $\mathbb{Z}$-structure in the Hecke algebra $\mathcal{H}_K$. It preserves the lattice $\mathcal{A}_{K,\mathbb{Z}}$ consisting of $\mathbb{Z}$-valued functions in $\mathcal{A}_K$. If L is a number field, we then define $\mathcal{H}_{K,L}$ and $\mathcal{A}_{K,L}$ by extension of scalars. We choose the field L so large that $\mathrm{Aut}(\mathbb{C}/L)$ fixes the simple $\mathcal{H}_K$-submodules of $\mathcal{A}_K$. Then

$$\Pi_f^K \simeq \mathbb{C} \otimes_L \Pi_f^K(L), \quad \Pi_f^K(L) = \Pi_f^K \cap \mathcal{A}_{K,L}.$$

Moreover, the L-model $\Pi_f^K(L)$ is unique up to complex scalars. We retain the decomposition of $\mathcal{A}_{K,L}$ into a sum of the various $\Pi_f^K(L)$, where Π runs through the usual set of automorphic representations with Π_∞ trivial and $\Pi_f^K \neq 0$.

2.1.3 *Integral models.* Let λ be a finite place of L above ℓ, and let $\mathcal{O}$ denote the ring of integers in the completion L_λ. Then, after extending scalars, $\mathcal{A}_{K,L_\lambda}$ is the sum of the $\Pi_f^K(L_\lambda)$ obtained from $\Pi_f^K(L)$ by tensoring with L_λ. These components have natural integral models obtained simply by intersecting with $\mathcal{A}_{K,\mathcal{O}}$; that is,

$$\Pi_f^K(L_\lambda) \simeq L_\lambda \otimes_{\mathcal{O}} \Pi_f^K(\mathcal{O}), \quad \Pi_f^K(\mathcal{O}) = \Pi_f^K(L_\lambda) \cap \mathcal{A}_{K,\mathcal{O}}.$$

However, the integral models $\Pi_f^K(\mathcal{O})$ need not be unique up to scalars. We also remark that their sum need not exhaust $\mathcal{A}_{K,\mathcal{O}}$, although the quotient is at worst torsion.

2.1.4 *Algebraic modular forms mod ℓ.* Let $\mathbb{F}$ be the residue field of $\mathcal{O}$. By the Brauer–Nesbitt principle (see [Vig96, p. 80]), the semisimplification is independent of the lattice (up to isomorphism) so that we have the usual decomposition, after semisimplification, in the form

$$\mathcal{A}_{K,\mathbb{F}}^{\mathrm{ss}} \simeq \bigoplus_{\Pi} m(\Pi)\Pi_f^K(\mathbb{F}),$$

where $\Pi_f^K(\mathbb{F})$ denotes the semisimplification of $\mathbb{F} \otimes_{\mathcal{O}} \Pi_f^K(\mathcal{O})$. Indeed, we consider the reductions of the lattice $\mathcal{A}_{K,\mathcal{O}}$ and the sublattice discussed above,

$$\bigoplus_{\Pi} m(\Pi)\Pi_f^K(\mathcal{O}),$$

and then take their semisimplifications. We say that K is a good small subgroup if the Hecke-module Π_f^K determines the representation Π_f up to isomorphism.

Lemma 2.1. *Suppose K is a good small subgroup. For ℓ outside a finite set of primes determined by K, the following hold: the $\mathcal{H}_{K,\mathbb{F}}$-module $\mathcal{A}_{K,\mathbb{F}}$ is semisimple, all the $\Pi_f^K(\mathbb{F})$ are simple, and $\Pi_f^K(\mathbb{F})$ occurs with multiplicity $m(\Pi)$.*

Proof. In this proof, let $H_{K,\mathbb{Z}}$ denote the image of $\mathcal{H}_{K,\mathbb{Z}}$ in $\mathrm{End}\mathcal{A}_{K,\mathbb{Z}}$. The algebra $H_{K,\mathbb{Z}}$ comes endowed with a natural symmetric pairing given by the trace of the composition. We consider its discriminant,

$$\det\{\mathrm{tr}\,(T_i \circ T_j)\} \in \mathbb{Z} - \{0\},$$

where $\{T_i\}$ is a basis for $H_{K,\mathbb{Z}}$. To see that this is non-zero, let H_K denote the algebra acting faithfully on $\mathcal{A}_K$. It is semisimple, so the natural pairing on H_K is non-degenerate since its radical is contained in the Jacobson radical. Now let ℓ be a prime not dividing the discriminant. Then the extended pairing on $\mathbb{F} \otimes_{\mathbb{Z}} H_{K,\mathbb{Z}}$ is non-degenerate, and this algebra is therefore semisimple. It follows that

$$\mathbb{F} \otimes_{\mathbb{Z}} H_{K,\mathbb{Z}} \twoheadrightarrow H_{K,\mathbb{F}}$$

is an isomorphism since its kernel is nilpotent. Consequently, $H_{K,\mathbb{F}}$ is semisimple. We now proceed to compute its dimension in two different ways. Decompose $\Pi_f^K(\mathbb{F})$ into simple submodules X

with multiplicity $m_\Pi(X)$. By Wedderburn theory, H_K is a product of matrix algebras. Hence, by computing $\dim_{\mathbb{C}} H_K = \dim_{\mathbb{F}} H_{K,\mathbb{F}}$ in two ways, we obtain the equality

$$\sum_X (\dim X)^2 = \sum_{\Pi,X} m_\Pi(X)^2 (\dim X)^2 + \text{mixed terms},$$

where the mixed terms are contributions coming from distinct types in $\Pi_f^K(\mathbb{F})$:

$$\sum_{\Pi, X\neq X'} m_\Pi(X) m_\Pi(X')(\dim X)(\dim X').$$

We deduce that there are no mixed terms, and each X is a constituent of a unique $\Pi_f^K(\mathbb{F})$ with multiplicity one. Therefore the $\Pi_f^K(\mathbb{F})$ must be simple and inequivalent as Π varies. □

If ℓ does not divide the order of the finite group Γ_i for all i, we can define a Hecke-compatible non-degenerate pairing on $\mathcal{A}_{K,\mathbb{F}}$ by the previous formula. This is automatic when ℓ is sufficiently large. More precisely, one has the following statement.

Lemma 2.2. *$\ell > 5 \Rightarrow \ell$ does not divide the $\#\Gamma_i$.*

Proof. Suppose $\ell > 5$ divides $\#\Gamma_i$. Then ℓ divides the pro-order of $\mathrm{GSp}(4, \mathbb{Z}_p)$,

$$\#\mathrm{GSp}(4, \mathbb{F}_p) = p^4(p-1)(p^2-1)(p^4-1),$$

for almost all p. Hence p has order at most 4 mod ℓ, contradicting Dirichlet. □

2.2 Generalized eigenspaces

Let π be a fixed automorphic representation of $G^{\mathrm{ad}}(\mathbb{A})$ such that π_∞ is trivial and $\pi_f^K \neq 0$. By Schur's lemma, the center $Z(\mathcal{H}_{K,\mathbb{Z}})$ acts on π_f^K by a character

$$\eta : Z(\mathcal{H}_{K,\mathbb{Z}}) \to L.$$

Here we may have to enlarge the field L. Since $Z(\mathcal{H}_{K,\mathbb{Z}})$ preserves $\mathcal{A}_{K,\mathbb{Z}}$, the values of η are in fact algebraic integers. We denote by $\bar{\eta}$ its reduction modulo λ and look at its generalized eigenspace

$$\mathcal{A}_{K,\mathbb{F}}(\bar{\eta}) = \{f \in \mathcal{A}_{K,\mathbb{F}} : \exists\, n \text{ such that } (r(\phi) - \bar{\eta}(\phi))^n f = 0 \ \forall \phi \in Z(\mathcal{H}_{K,\mathbb{F}})\}.$$

This subspace is preserved by $\mathcal{H}_{K,\mathbb{F}}$. Its semisimplification is given as follows.

Lemma 2.3. *$\mathcal{A}_{K,\mathbb{F}}(\bar{\eta})^{ss} \simeq \bigoplus_{\Pi:\Pi\equiv\pi \ (\mathrm{mod}\ \lambda)} m(\Pi)\Pi_f^K(\mathbb{F})$.*

Proof. In this proof, $\mathbb{T}$ denotes the image of $Z(\mathcal{H}_{K,\mathcal{O}})$ in $\mathrm{End}\mathcal{A}_{K,\mathcal{O}}$. Since $\mathcal{O}$ is complete, $\mathbb{T}$ is the direct product of its localizations. Clearly, $\bar{\eta}$ factors through $\mathbb{T}$, and we let $\mathfrak{m} = \ker(\bar{\eta})$ be the corresponding maximal ideal in $\mathbb{T}$. Then $(\mathcal{A}_{K,\mathcal{O}})_{\mathfrak{m}} = \mathbb{T}_{\mathfrak{m}} \otimes_{\mathbb{T}} \mathcal{A}_{K,\mathcal{O}}$ is a lattice in

$$(\mathcal{A}_{K,L_\lambda})_{\mathfrak{m}} = \bigoplus_{\eta'\equiv\eta \ (\mathrm{mod}\ \lambda)} \mathcal{A}_{K,L_\lambda}(\eta') \simeq \bigoplus_{\Pi:\Pi\equiv\pi \ (\mathrm{mod}\ \lambda)} m(\Pi)\Pi_f^K(L_\lambda).$$

Also, it is clearly true that $(\mathcal{A}_{K,\mathbb{F}})_{\mathfrak{m}} = \mathcal{A}_{K,\mathbb{F}}(\bar{\eta})$. Now invoke the Brauer–Nesbitt principle. □

We will assume that $\ell > 5$ from now on.

Lemma 2.4. *$\mathcal{A}_{K,\mathbb{F}}(\bar{\eta})$ is self-dual as a module over $\mathcal{H}_{K,\mathbb{F}}$.*

Proof. From an easy inductive argument based on the socle filtration, it follows that the pairing on $\mathcal{A}_{K,\mathbb{F}}$ restricts to a non-degenerate pairing between the generalized eigenspaces $\mathcal{A}_{K,\mathbb{F}}(\bar{\eta})$ and $\mathcal{A}_{K,\mathbb{F}}(\bar{\eta}^\vee)$. However, it is well-known that $\pi \simeq \pi^\vee$ (this is even true locally for any odd-rank special orthogonal group), so $\bar{\eta} = \bar{\eta}^\vee$. The Hecke actions are intertwined by the compatibility relation. □

2.3 The universal module

In this section we fix a prime $q \neq r$. We let $G = G(\mathbb{Q}_q)$ and fix a hyperspecial subgroup K. Also, we fix a Borel subgroup B. Let $\mathbb{F}$ be an algebraic closure of $\mathbb{F}_\ell$ and consider the spherical Hecke algebra $\mathcal{H}_{K,\mathbb{F}}$. We look at the degree character, giving the Hecke-action on the trivial representation

$$\deg : \mathcal{H}_{K,\mathbb{F}} \to \mathbb{F}.$$

We define a category with objects (V, v), where V is a smooth G-representation over $\mathbb{F}$ and $v \in V$ is a K-fixed vector on which $\mathcal{H}_{K,\mathbb{F}}$ acts by deg, and with the obvious morphisms. This category has a universal initial object. An explicit construction realizes it as the induced module

$$\mathcal{M} = C_c(G/K) \otimes_{\mathcal{H}_{K,\mathbb{F}}} \mathbb{F}_{\deg}.$$

Obviously, $\mathcal{M}^K$ is spanned by the class of e_K, the neutral element in $\mathcal{H}_{K,\mathbb{F}}$. Also observe that $\mathcal{M}$ is generated by its K-invariants and hence is cyclic. We will need the following theorem of Lazarus [Laz00].

THEOREM 2.5. *Suppose $q \neq \ell$ and $q^4 \neq 1 \pmod{\ell}$. Then $\mathcal{M}^\vee \simeq C^\infty(B\backslash G)$.*

Proof. Let δ_B denote the modulus character of B. Note that $C^\infty(B\backslash G)^\vee$ is none other than the principal series of $\delta_B^{1/2}$. By the universal property of $\mathcal{M}$, there is a canonical surjective G-map

$$\mathcal{M} \to C^\infty(B\backslash G)^\vee.$$

By assumption, ℓ is banal for q, that is, $\ell \neq q$ does not divide $\#G(\mathbb{F}_q)$. Therefore, [Laz00, Theorem 1.0.3] applies. Consequently, the map must be an isomorphism since the two representations have the same semisimplifications. □

We say that ℓ is banal for q if it satisfies the hypothesis of this theorem. Note that we must then have $\ell > 5$. The result allows us to write down a composition series for $\mathcal{M}^\vee$. There are two parabolic subgroups containing B: the Klingen parabolic P_α and the Siegel parabolic P_β. The latter has abelian unipotent radical. Let us take B to be the subgroup of upper triangular matrices:

$$B = \left\{ \begin{pmatrix} a & & & \\ & b & & \\ & & cb^{-1} & \\ & & & ca^{-1} \end{pmatrix} \begin{pmatrix} 1 & x & & \\ & 1 & & \\ & & 1 & -x \\ & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & r & s \\ & 1 & t & r \\ & & 1 & \\ & & & 1 \end{pmatrix} \right\}.$$

Then, the two maximal parabolic subgroups have the following realizations:

$$P_\alpha = \left\{ \begin{pmatrix} c & & \\ & g & \\ & & c^{-1}\det g \end{pmatrix} \begin{pmatrix} 1 & x & & \\ & 1 & & \\ & & 1 & -x \\ & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & r & s \\ & 1 & & r \\ & & 1 & \\ & & & 1 \end{pmatrix} \right\}$$

and

$$P_\beta = \left\{ \begin{pmatrix} g & \\ & c\,{}^\tau g^{-1} \end{pmatrix} \begin{pmatrix} 1 & & r & s \\ & 1 & t & r \\ & & 1 & \\ & & & 1 \end{pmatrix} \right\},$$

where ${}^\tau g$ denotes the skew-transpose of g, i.e. the transpose with respect to the second diagonal. Now suppose that $q^i \neq 1 \pmod{\ell}$ for $i = 1, \ldots, 4$. Then, by [Laz00, Theorem 4.7.2], the natural filtration,

$$0 \subset \{\text{constants}\} \subset C^\infty(P_\alpha \backslash G) \subset C^\infty(P_\alpha \backslash G) + C^\infty(P_\beta \backslash G) \subset C^\infty(B \backslash G),$$

has subquotients $\mathbf{1}$, V_α, V_β and St, all occurring with multiplicity one.

2.4 Existence of certain subquotients

Let π be as above, but assume that $m(\pi) = 1$. Moreover, suppose K is a good small subgroup and that ℓ lies outside a finite set of primes as in Lemma 2.1.

Let $q \notin \{\ell, r\}$ be a prime such that:

- $K = K_q K^q$ with K_q hyperspecial;
- $q^i \neq 1 \pmod{\ell}$ for $i = 1, \ldots, 4$;
- $\pi_q \equiv \mathbf{1} \pmod{\lambda}$.

Then fix an Iwahori subgroup $I_q \subset K_q$ and let $I = I_q K^q$. By our assumptions on π and ℓ, we can identify $\pi_f^K(\mathbb{F})$ with a submodule of $\mathcal{A}_{K,\mathbb{F}}(\bar\eta)$. We look at the Iwahori-modules they generate, that is, the submodules

$$\mathcal{H}_{I,\mathbb{F}} \cdot \pi_f^K(\mathbb{F}) \subset \mathcal{H}_{I,\mathbb{F}} \cdot \mathcal{A}_{K,\mathbb{F}}(\bar\eta) \subset \mathcal{A}_{I,\mathbb{F}}(\bar\eta).$$

Here we abuse notation a bit and let $\mathcal{A}_{I,\mathbb{F}}(\bar\eta)$ denote the generalized eigenspace of $\bar\eta$ composed with the canonical homomorphism

$$Z(\mathcal{H}_{I,\mathbb{F}}) \to Z(\mathcal{H}_{K,\mathbb{F}}), \quad \phi \mapsto e_K \star \phi.$$

The connection with the universal module $\mathcal{M}$ is given by the next lemma.

Lemma 2.6. *The module $\mathcal{M}$ has a quotient $\mathcal{N}$ such that, as modules over $\mathcal{H}_{I,\mathbb{F}}$,*

$$\mathcal{N}^I \otimes \pi_f^K(\mathbb{F}) \simeq \mathcal{H}_{I,\mathbb{F}} \cdot \pi_f^K(\mathbb{F}).$$

Moreover, if π_q is a full unramified principal series, we have $\mathcal{M} = \mathcal{N}$.

Proof. Note that $\mathcal{H}_{I,\mathbb{F}} \cdot \pi_f^K(\mathbb{F})$ is a multiple of $\pi_f^K(\mathbb{F})$, viewed as a simple module over $\mathcal{H}_{K^q,\mathbb{F}}$. By [Vig96, Theorem 3.12], since ℓ is banal for q, there is a representation $\mathcal{N}$ of $G(\mathbb{Q}_q)$ such that

$$\mathcal{N}^I \simeq \mathrm{Hom}_{\mathcal{H}_{K^q,\mathbb{F}}}(\pi_f^K(\mathbb{F}), \mathcal{H}_{I,\mathbb{F}} \cdot \pi_f^K(\mathbb{F})).$$

Moreover, $\mathcal{N}$ is generated by its Iwahori-invariants. It remains to show that $\mathcal{N}$ is a quotient of $\mathcal{M}$. By the universal property of $\mathcal{M}$, there is a canonical surjective map of $\mathcal{H}_{I,\mathbb{F}}$-modules,

$$\mathcal{M}^I \otimes \pi_f^K(\mathbb{F}) \twoheadrightarrow \mathcal{H}_{I,\mathbb{F}} \cdot \pi_f^K(\mathbb{F}).$$

This, in turn, defines a surjective map $\mathcal{M}^I \twoheadrightarrow \mathcal{N}^I$; indeed, it is enough to show surjectivity after tensoring with $\pi_f^K(\mathbb{F})$. By the result of Vigneras mentioned above, this map comes from a map of representations $\mathcal{M} \to \mathcal{N}$, which must be surjective since $\mathcal{N}^I$ generates $\mathcal{N}$. Now let us assume that π_q is generic and show that the above canonical map is injective; we do this by comparing dimensions. Obviously, the source has dimension $8\dim \pi_f^K$. Furthermore,

$$\mathcal{H}_{I,\mathbb{F}} \cdot \pi_f^K(\mathbb{F}) \simeq \mathbb{F} \otimes_{\mathcal{O}} \pi_f^I(\mathcal{O}).$$

As π_q is generic, $\dim \pi_q^{I_q} = 8$, hence $\mathcal{H}_{I,\mathbb{F}} \cdot \pi_f^K(\mathbb{F})$ has dimension $8\dim \pi_f^K$. □

LEMMA 2.7. *$\mathcal{N}$ is the trivial representation only if π is abelian modulo λ.*

Proof. Suppose $\mathcal{N} = \mathbf{1}$. Then $\mathcal{A}_{I,\mathbb{F}}(\bar{\eta})$ contains an eigenform f such that the Iwahori–Hecke algebra acts on $\mathbb{F} \cdot f$ by the degree character. Therefore, f is $G(\mathbb{Q}_q)$-invariant (on both sides). Note that

$$1 \to G^{\mathrm{der}}(\mathbb{A}_f) \to G(\mathbb{A}_f) \xrightarrow{c} \mathbb{A}_f^* \to 1$$

is exact since $H^1(\mathbb{Q}_p, G^{\mathrm{der}}) = 1$, as G^{der} is simple and simply connected. We claim that f factors through c. This follows easily from strong approximation, using the fact that $G^{\mathrm{der}}(\mathbb{Q}_q)$ is non-compact. Thus $\bar{\eta}$ occurs in the space of $\mathbb{F}$-valued functions on the finite abelian group $\mathbb{A}_f^*/\mathbb{Q}_+^* c(I)$. By the Deligne–Serre lifting lemma [DS74, Lemma 6.11], there is a character $\eta' \equiv \eta \pmod{\lambda}$ occurring in the space $\mathbb{C}$-valued functions. The group characters form a basis for this space, so there is an automorphic character χ of $G(\mathbb{A}_f)$ such that $\eta' = \eta_\chi$. □

In what follows, we use the notation from [Sch05]; see Appendix A and Appendix B.

LEMMA 2.8. *Assume that π is non-abelian modulo λ. Then π_q must be of type I, IIb or IIIb. Accordingly, we have the following three possibilities for $\mathcal{N}^\vee$:*

$$\mathcal{N}^\vee \simeq C^\infty(P\backslash G) \quad \textit{where} \quad P = \begin{cases} B & \textit{if } \pi_q \textit{ is of type I,} \\ P_\alpha & \textit{if } \pi_q \textit{ is of type IIb,} \\ P_\beta & \textit{if } \pi_q \textit{ is of type IIIb.} \end{cases}$$

Proof. The trivial representation is the unique irreducible quotient of $\mathcal{M}$, so it is also a quotient of $\mathcal{N}$. However, $\mathcal{N} \neq \mathbf{1}$ by the previous lemma. Write down a composition series for $\mathcal{N}^\vee$ of the form

$$0 \overset{\mathbf{1}}{\subset} \mathbf{1} \overset{V}{\subset} X \subset \cdots \subset \mathcal{N}^\vee,$$

with irreducible subquotients. Here X is a non-trivial extension of V by $\mathbf{1}$. Otherwise $V^\vee$ is a quotient of $\mathcal{M}$ and hence trivial. However, all constituents of $\mathcal{M}$ occur with multiplicity one. Thus

$$\mathrm{Ext}^1(\mathbf{1}, V) \neq 0$$

by self-duality. According to [Clo00], the arguments in Casselman's paper [Cas81] remain valid as long as ℓ is banal for q. Therefore, $V \simeq V_P$ for a maximal parabolic subgroup P. Moreover, there is an isomorphism

$$X \simeq C^\infty(P\backslash G).$$

Suppose $P = P_\alpha$. Then $C^\infty(P_\alpha\backslash G)^\vee$ is a quotient of $\mathcal{N}$. In turn, there is a map

$$\mathcal{N}^J \otimes \pi_f^K(\mathbb{F}) \simeq \mathcal{H}_{J,\mathbb{F}} \cdot \pi_f^K(\mathbb{F}) \twoheadrightarrow C^\infty(P_\alpha\backslash G/J)^\vee \otimes \pi_f^K(\mathbb{F})$$

for any J. If we take $J = J_\beta$, we deduce that $\dim \pi_q^{J_\beta}$ is at least 3. Since π_q is also unramified, it follows from Appendix B that it must be of type I or IIb. In the latter case, note that $\dim \mathcal{N}^I$ and $\#P_\alpha\backslash G/I$ both equal 4. Similarly, if $P = P_\beta$, we conclude that π_q must be of type I or IIIb. □

2.5 Proof of Theorem A

Let P be a maximal parabolic subgroup such that $\mathcal{N}^\vee$ contains $C^\infty(P\backslash G)$. Thus $P=P_\alpha$ if π_q is of type IIb, and $P=P_\beta$ if π_q is of type IIIb. When π_q is generic, P can be taken to be arbitrary.

Lemma 2.9. *The modules $V_P^I\otimes\pi_f^K(\mathbb{F})$ and $\mathbf{1}^I\otimes\pi_f^K(\mathbb{F})$ both occur with multiplicity one in $\mathcal{H}_{I,\mathbb{F}}\cdot\mathcal{A}_{K,\mathbb{F}}(\bar\eta)$.*

Proof. By the universality of $\mathcal{M}$, there is a canonical surjective $\mathcal{H}_{I,\mathbb{F}}$-map,

$$\mathcal{M}^I\otimes\mathcal{A}_{K,\mathbb{F}}(\bar\eta)\twoheadrightarrow\mathcal{H}_{I,\mathbb{F}}\cdot\mathcal{A}_{K,\mathbb{F}}(\bar\eta).$$

Recall that $\mathcal{M}$ satisfies multiplicity one, and $\pi_f^K(\mathbb{F})$ occurs once in $\mathcal{A}_{K,\mathbb{F}}(\bar\eta)$. □

Lemma 2.10. *$V_P^I\otimes\pi_f^K(\mathbb{F})$ occurs in the quotient $\mathcal{A}_{I,\mathbb{F}}(\bar\eta)/\mathcal{H}_{I,\mathbb{F}}\cdot\mathcal{A}_{K,\mathbb{F}}(\bar\eta)$.*

Proof. First we show that $V_P^I\otimes\pi_f^K(\mathbb{F})$ or $\mathbf{1}^I\otimes\pi_f^K(\mathbb{F})$ occurs in the quotient in the lemma; then we rule out the latter. Otherwise, both modules would occur with multiplicity one in $\mathcal{A}_{I,\mathbb{F}}(\bar\eta)$ by the previous lemma. Now, $\mathcal{A}_{I,\mathbb{F}}(\bar\eta)$ has a composition series where the constituent $V_P^I\otimes\pi_f^K(\mathbb{F})$ is the left neighbor of $\mathbf{1}^I\otimes\pi_f^K(\mathbb{F})$. Now we recall that $\mathcal{A}_{I,\mathbb{F}}(\bar\eta)$ is self-dual by Lemma 2.4.

Therefore,

$$C^\infty(P\backslash G/I)\otimes\pi_f^K(\mathbb{F})$$

is also a subquotient. In particular, it has a composition series where the constituents form a subseries of the above composition series. By multiplicity one, we must have an exact sequence

$$0\to V_P^I\otimes\pi_f^K(\mathbb{F})\to C^\infty(P\backslash G/I)\otimes\pi_f^K(\mathbb{F})\to\mathbf{1}^I\otimes\pi_f^K(\mathbb{F})\to 0.$$

However, this is impossible since $\mathbf{1}$ is not a quotient of $C^\infty(P\backslash G)$. Suppose $\mathbf{1}^I\otimes\pi_f^K(\mathbb{F})$ occurs in the quotient. Then there exists an automorphic representation $\Pi\equiv\pi\pmod{\lambda}$, with $\Pi_f^K=0$, such that $\mathbf{1}^I\otimes\pi_f^K(\mathbb{F})$ is a summand of $\Pi_f^I(\mathbb{F})$. Applying the idempotent e_K, we reach a contradiction. □

We can now finish the proof of Theorem A as follows. Suppose that $P=P_\alpha$. Then there exists an automorphic representation $\tilde\pi\equiv\pi\pmod{\lambda}$, with $\tilde\pi_f^K=0$, such that $V_\alpha^I\otimes\pi_f^K(\mathbb{F})$ is a summand of $\tilde\pi_f^I(\mathbb{F})$. Applying the idempotent e_{J_β}, we see from Appendix B that $\dim\tilde\pi_q^{J_\beta}$ is at least 2. Since $\tilde\pi_q$ is also ramified, we conclude (again using Appendix B) that it must be of type IIIa. The type IVb is immediately ruled out as it is not unitary. Analogously, if $P=P_\beta$, we deduce that $\tilde\pi_q$ is of type IIa.

3. Matching orbital integrals

3.1 Twisted orbital integrals

For the time being, we let G denote the non-split inner form of G' over $\mathbb{Q}_r$. It splits over the unramified quadratic extension E. Let θ be the generator of $\mathrm{Gal}(E/\mathbb{Q}_r)$, and fix an inner twisting ψ defined over E. We define a norm mapping N from $G(E)$ to itself by setting $N\delta=\delta\theta(\delta)$. For $\delta\in G(E)$, we then define $\mathcal{N}\delta$ by intersecting the stable conjugacy class of $N\delta$ with $G(\mathbb{Q}_r)$. It may happen that $\mathcal{N}\delta$ is empty, since G is not quasi-split. Otherwise, the stable twisted conjugacy class of δ is defined to be the fiber of $\mathcal{N}$ through δ; it is a finite union of twisted conjugacy classes. We consider the $\mathbb{Q}_r$-group I obtained from G by restriction of scalars from E. Then θ defines an automorphism of I over $\mathbb{Q}_r$, again denoted by θ. Now let $\delta\in G(E)$ be an element such that $N\delta$ is semisimple. The extended twisted centralizer of δ is the $\mathbb{Q}_r$-group $I_{\delta\theta}$ with rational points

$$I_{\delta\theta}(\mathbb{Q}_r) = \{x \in G(E) : x^{-1}\delta\theta(x) \in Z(\mathbb{Q}_r)\delta\}.$$

Choose measures on $G(E)$ and $I_{\delta\theta}(\mathbb{Q}_r)$, and consider the twisted orbital integral

$$O_{\delta\theta}(f_E) = \int_{I_{\delta\theta}(\mathbb{Q}_r)\backslash G(E)} f_E(x^{-1}\delta\theta(x))\, dx$$

of a function $f_E \in C_c^\infty(G^{\mathrm{ad}}(E))$. Now let $\{\tilde{\delta}\}$ be a set of representatives for the twisted conjugacy classes within the stable twisted conjugacy class of δ mod $Z(\mathbb{Q}_r)$. Then $I_{\tilde{\delta}\theta}$ is an inner form of $I_{\delta\theta}$, and we transform the measure as usual. Next, define the stable twisted orbital integral of f_E to be the sum

$$SO_{\delta\theta}(f_E) = \sum_{\tilde{\delta}} e(I_{\tilde{\delta}\theta}) O_{\tilde{\delta}\theta}(f_E).$$

To be precise, we put $SO_{\delta\theta}(f_E) = 0$ if $\mathcal{N}\delta$ is empty. Consider a test function $f \in C_c^\infty(G^{\mathrm{ad}}(\mathbb{Q}_r))$. We say that the two functions f and f_E have matching orbital integrals if for all semisimple $\gamma \in G(\mathbb{Q}_r)$,

$$SO_\gamma(f) = \begin{cases} SO_{\delta\theta}(f_E) & \text{if } \gamma \text{ belongs to } \mathcal{N}\delta \text{ mod } Z(\mathbb{Q}_r), \\ 0 & \text{if } \gamma \text{ does not come from } G(E). \end{cases}$$

We note that G_γ and $I_{\delta\theta}$ are inner forms if $\gamma \in \mathcal{N}\delta$, and we use compatible Haar measures on both sides. Also, the measures on $G(\mathbb{Q}_r)$ and $G(E)$ are fixed; in practice, they will be normalized compatibly. Finally, we note that all these definitions carry over to G'. Indeed, things are more well-behaved since G' is quasi-split: for example, the norm map $\mathcal{N}$ is defined everywhere.

3.2 Base change for idempotents

Let K be a paramodular subgroup of $G(\mathbb{Q}_r)$, and let x be the vertex in the tree fixed by K. Since $E/\mathbb{Q}_r$ is unramified, we may view x as a θ-invariant point in the building of G over E. Then let K_E be the parahoric subgroup of $G(E)$ fixing x. We choose measures on $G(\mathbb{Q}_r)$ and $G(E)$ such that K and K_E have the same measure. The following crucial result is due to Kottwitz [Kot86].

THEOREM 3.1. *The idempotents e_K and e_{K_E} have matching orbital integrals.*

Proof. Let L be the completion of the maximal unramified extension of $\mathbb{Q}_r$, and let σ be the Frobenius over $\mathbb{Q}_r$. We view x as a point in the building of G over L and let K_L be the open bounded subgroup of $G(L)$ fixing x. We claim that K_L satisfies the three conditions on [Kot86, p. 240]. Clearly, K_L is fixed by σ. Secondly, to see that $k \mapsto k^{-1}\sigma(k)$ defines a surjective map from K_L to itself, we argue as in [Kot86]. Specifically, let $\mathcal{G}$ be the smooth affine group scheme over $\mathbb{Z}_r$ attached to x in Bruhat–Tits theory; it has generic fiber G and $\mathcal{O}_L$-points K_L. Since G^{der} is simply connected, the special fiber $\bar{\mathcal{G}}$ is connected and we can refer to [Gre63, Proposition 3]. Now the result follows by paraphrasing the arguments in [Kot86] with our definition of orbital integrals. □

Similarly, we let K' be the paramodular subgroup of $G'(\mathbb{Q}_r)$ fixing the vertex x' in the building. Extending scalars, we view x' as a point in the building of G' over E, and we let K'_E be the parahoric subgroup of $G'(E)$ fixing x'.

THEOREM 3.2. *The functions $e_{\eta K'}$ and $e_{\eta K'_E}$ have matching orbital integrals.*

Proof. As before, the idempotents $e_{K'}$ and $e_{K'_E}$ have matching orbital integrals. We have

$$e_{\langle\eta,K'\rangle} = e_{K'} + e_{\eta K'},$$

and similarly over E, so it remains to show that $e_{\langle\eta,K'\rangle}$ and $e_{\langle\eta,K'_E\rangle}$ match. Again, this follows from the arguments in [Kot86]. However, the proof is not as straightforward as above, since the group $\langle\eta, K'_L\rangle$ does not satisfy the conditions on [Kot86, p. 240]. Indeed, the map $k \mapsto k^{-1}\sigma(k)$ only maps onto the subgroup K'_L. This, however, is sufficient: in fact, using the notation of [Kot86], it is still true that X mod center is identified with the set of fixed points of the Frobenius σ on X_L modulo the center. □

3.3 The comparison over *E*

It is well-known that G' has a unique inner form over $\mathbb{Q}_r$. Thus, by the inflation–restriction sequence, we compute the Galois cohomology of the adjoint group:

$$H^1(E/\mathbb{Q}_r, G'^{\mathrm{ad}}(E)) \simeq H^1(\mathbb{Q}_r, G'^{\mathrm{ad}}) \simeq \{\pm 1\}.$$

The non-trivial cohomology class is represented by the cocycle $\theta \mapsto \eta$. We may therefore choose our twisting ψ such that $\theta\psi \circ \psi^{-1}$ is conjugation by η. With this choice of ψ, the following integrals match.

Lemma 3.3. *$O_{\delta\theta}(e_{K_E}) = O_{\delta'\theta}(e_{\eta K'_E})$ where $\delta' = \psi(\delta)\eta^{-1}$.*

Proof. Obviously, ψ restricts to an isomorphism between $I_{\delta\theta}$ and $I_{\delta'\theta}$. Moreover, a simple explicit computation shows that

$$O_{\delta\theta}(e_{K_E}) = O_{\delta'\theta}(e_{\psi(K_E)\eta^{-1}}).$$

Now, $\psi^{-1}(K'_E)$ is θ-invariant, hence it stabilizes a conjugate of x so that

$$\psi(K_E) = \xi K'_E \xi^{-1}$$

for some $\xi \in G'(E)$. It follows that $\theta(\xi)^{-1}\eta\xi$ normalizes K'_E, and then $\psi(K_E)\eta^{-1}$ is a θ-conjugate of $\eta K'_E$ mod center. Their characteristic functions therefore have the same twisted orbital integrals. □

Lemma 3.4. *$SO_{\delta\theta}(e_{K_E}) = SO_{\delta'\theta}(e_{\eta K'_E})$ where $\delta' = \psi(\delta)\eta^{-1}$.*

Proof. First we deal with the case where $\mathcal{N}\delta$ is non-empty. Let $\{\tilde{\delta}\}$ be a set of representatives for the twisted conjugacy classes within the stable twisted conjugacy class of δ mod $Z(\mathbb{Q}_r)$. It is straightforward to check that $\{\tilde{\delta}'\}$ is then an analogous set of representatives for δ'. The result then follows from the previous lemma. If $\mathcal{N}\delta$ is empty, it remains to show the vanishing statement

$$SO_{\delta'\theta}(e_{\eta K'_E}) = 0.$$

Suppose otherwise; then $O_{\tilde{\delta}'\theta}(e_{\eta K'_E})$ is non-zero for some $\tilde{\delta}'$. However, it equals $O_{\tilde{\delta}\theta}(e_{K_E})$ by the previous lemma. By the theorem on [Kot86, p. 243], there is a corresponding $\tilde{\gamma} \in \mathcal{N}\tilde{\delta}$. Hence $\mathcal{N}\delta$ is non-empty, and this is a contradiction. □

3.4 Proof of Theorem B

To prove Theorem B, let $\gamma' \in G'(\mathbb{Q}_r)$ be an arbitrary semisimple element. First, we assume that γ' does not come from $G(\mathbb{Q}_r)$. Then, we must show that $SO_{\gamma'}(e_{\eta K'})$ vanishes. We may clearly assume that γ' belongs to $\mathcal{N}\delta'$ for some $\delta' \in G'(E)$. Write $\delta' = \psi(\delta)\eta^{-1}$; then, from what we have shown,

$$SO_{\gamma'}(e_{\eta K'}) = SO_{\delta'\theta}(e_{\eta K'_E}) = SO_{\delta\theta}(e_{K_E}).$$

As a result, it suffices to show that $\mathcal{N}\delta$ is empty. If $\mathcal{N}\delta$ is non-empty, there must exist a $\gamma \in G(\mathbb{Q}_r)$ that is stably conjugate to $N\delta$ mod center. However, $\psi(N\delta) = rN\delta'$, so $\psi(\gamma)$ would then be stably conjugate to γ' modulo the center; but this contradicts our assumption that γ' does not come from $G(\mathbb{Q}_r)$. Next, we assume that γ' is stably conjugate to $\psi(\gamma)$ for some $\gamma \in G(\mathbb{Q}_r)$. We must show that

$$SO_{\gamma'}(e_{\eta K'}) = SO_{\gamma}(e_K).$$

It is easy to check that $\gamma \in \mathcal{N}\delta$ mod center if and only if $\gamma' \in \mathcal{N}\delta'$ mod center. If this does not hold, then both sides are zero. If it does hold, the first string of equalities can be extended by $SO_{\gamma}(e_K)$. □

4. Functoriality

4.1 Endoscopy

4.1.1 *The endoscopic group H.* Up to equivalence, G, or its class of inner forms, admits a unique non-trivial elliptic endoscopic triple (H, s, ξ). The underlying endoscopic group is

$$H = (\mathrm{GL}(2) \times \mathrm{GL}(2))/\mathbb{G}_m,$$

where $\mathbb{G}_m$ is embedded in the center by identifying x with (x, x^{-1}). The dual is

$$\hat{H} = \{(x, x') \in \mathrm{GL}(2,\mathbb{C}) \times \mathrm{GL}(2,\mathbb{C}) : \det x = \det x'\}.$$

There is a natural embedding $\xi : \hat{H} \hookrightarrow \hat{G}$ defined as follows:

$$\xi : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \times \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \mapsto \begin{pmatrix} a' & & & b' \\ & a & b & \\ & c & d & \\ c' & & & d' \end{pmatrix}.$$

In this way, $\hat{H}$ is being identified with the centralizer of the semisimple element

$$s = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix}.$$

4.1.2 *Transfer and the fundamental lemma.* Let p be a prime. A semisimple element $\delta \in H(\mathbb{Q}_p)$ is said to be (G, H)-regular if $\alpha(\delta) \neq 1$ for every root α of G that does not come from H. We have the following fundamental result in our case.

Theorem 4.1. *For every test function $f \in C_c^\infty(G^{\mathrm{ad}}(\mathbb{Q}_p))$ there exists a matching function $f^H \in C_c^\infty(H^{\mathrm{ad}}(\mathbb{Q}_p))$, that is, a function such that*

$$SO_\delta(f^H) = \sum_\gamma \Delta_{G,H}(\delta, \gamma) e(G_\gamma) O_\gamma(f)$$

for all (G, H)-regular semisimple $\delta \in H(\mathbb{Q}_p)$. Here the sum runs over a set of representatives for the conjugacy classes in $G(\mathbb{Q}_p)$ belonging to the stable conjugacy class associated to δ. We use compatible measures on both sides. The $\Delta_{G,H}(\delta, \gamma)$ are the Langlands–Shelstad transfer factors.

Proof. By a descent argument due to Langlands and Shelstad (see [LS90, p. 495]), it suffices to prove the theorem for G and its centralizers near the identity. Here we have Shalika germ expansions of the orbital integrals, and Hales computed and matched these germs in [Hal89]. □

We remark that since H has no endoscopy itself, SO_δ equals O_δ up to a sign. We also note that, by a more recent result of Waldspurger (see [Wal97, p. 157]), the previous theorem in fact follows from the following supplementary result known as the standard fundamental lemma.

THEOREM 4.2. *Let $p \neq r$, and let K and K_H be hyperspecial subgroups of $G(\mathbb{Q}_p)$ and $H(\mathbb{Q}_p)$, respectively. Then, if f equals the characteristic function e_K, we may take f^H above to be the characteristic function e_{K_H} (up to a constant).*

Proof. This is due to Hales [Hal97]. □

By [Wal97], one can also transfer functions f on $G(\mathbb{Q}_p)$ to functions $f^{G'}$ on $G'(\mathbb{Q}_p)$ with matching orbital integrals. The archimedean case of Theorem 4.1 was proved by Shelstad in [She79] and [She82]. Finally, we mention that, of course, all we have said is also true for G'.

4.1.3 *Local character identities.* Let ρ be an irreducible admissible representation of $H(\mathbb{Q}_p)$. It factors as $\rho_1 \otimes \rho_2$, where the ρ_i are representations of GL(2) with the same central character. Since H has no endoscopy, the character tr ρ is a stable distribution. By results of Arthur [Art96] and Shelstad [She82], there is an expansion

$$\mathrm{tr}\,\rho(f^H) = \sum_{\pi} \Delta_{G,H}(\rho,\pi)\mathrm{tr}\,\pi(f)$$

for any $f \in C_c^\infty(G(\mathbb{Q}_p))$. Here π runs over irreducible representations of $G(\mathbb{Q}_p)$, and the $\Delta_{G,H}(\rho,\pi)$ are spectral analogues of the Langlands–Shelstad transfer factors. There is a similar expansion of tr ρ in terms of representations of $G'(\mathbb{Q}_p)$. Using θ-correspondence, Weissauer has made this expansion explicit in [Wei1] and [Wei2]. We recall his results below. There are precisely two isomorphism classes of quaternary quadratic spaces X with discriminant one, namely the split space X^s and the anisotropic space X^a. Now, the key is the two identifications

$$\mathrm{GSO}(X^s) \simeq H, \quad \mathrm{GSO}(X^a) \simeq \breve{H} = (D^* \times D^*)/\mathbb{G}_m.$$

Here D is the division quaternion algebra over $\mathbb{Q}_p$, and GSO(X) denotes the identity component of the orthogonal similitude group GO(X). Note that, by the Jacquet–Langlands correspondence, there is a one-to-one correspondence between irreducible representations $\breve{\rho}$ of $\breve{H}(\mathbb{Q}_p)$ and discrete series representations ρ of $H(\mathbb{Q}_p)$. Later, we will transfer representations to GSp(4) by using θ-correspondence for similitude groups. For that purpose, we now briefly review a result of Roberts [Rob96] on θ-correspondence in our special case. Assume that p is odd, and fix a non-trivial character of $\mathbb{Q}_p$. Correspondingly, we have the Weil representation ω of $\mathrm{Sp}(4) \times \mathrm{O}(X)$ on the Schwartz space of X^2. It extends naturally, with formulas as given in [Rob96], to a representation $\tilde{\omega}$ of the group

$$\{(x,x') \in \mathrm{GSp}(4) \times \mathrm{GO}(X) : c(x) = c(x')\}.$$

Let $\mathcal{R}_X(\mathrm{GSp}(4))$ denote the set of irreducible representations π of GSp(4) such that the restriction to Sp(4) is multiplicity-free and has a constituent that is a quotient of the usual Weil representation ω. Define the set $\mathcal{R}_4(\mathrm{GO}(X))$ similarly. Since disc $X = 1$, the condition $\mathrm{Hom}(\tilde{\omega}, \pi \otimes \rho) \neq 0$ defines a bijection

$$\mathcal{R}_X(\mathrm{GSp}(4)) \leftrightarrow \mathcal{R}_4(\mathrm{GO}(X))$$

by [Rob96]. We denote this map and its inverse by $\pi \mapsto \theta(\pi)$ and $\rho \mapsto \theta(\rho)$, respectively. These maps turn out to be *independent* of the additive character used to define ω. Now let $\rho = \rho_1 \otimes \rho_2$ be a representation of $\mathrm{GSO}(X)$; it is said to be regular if $\rho_1 \neq \rho_2$. In this case, by Mackey theory, the induced representation of ρ to $\mathrm{GO}(X)$ is irreducible and we denote it by ρ^+. When $\rho_1 = \rho_2$, we say that ρ is invariant. If so, it has exactly two extensions to $\mathrm{GO}(X)$. However, by [Rob99], there is a *unique* extension ρ^+ occurring in the θ-correspondence with $\mathrm{GSp}(4)$.

THEOREM 4.3. *Let ρ be a discrete series representation of $H(\mathbb{Q}_p)$. Then we have the endoscopic character relation*

$$\operatorname{tr} \rho(f^H) = \operatorname{tr} \theta(\rho^+)(f) - \operatorname{tr} \theta(\breve{\rho}^+)(f)$$

for any $f \in C_c^\infty(G(\mathbb{Q}_p))$. Here $\breve{\rho}$ is the Jacquet–Langlands correspondent of ρ.

Proof. This is due to Weissauer; see [Wei2, Proposition 1]. □

Weissauer makes the following supplementary remarks. The lift $\theta(\rho^+)$ is always generic, whereas $\theta(\breve{\rho}^+)$ is non-generic. If ρ is regular, both θ-lifts are discrete series representations (indeed, $\theta(\breve{\rho}^+)$ is always supercuspidal). On the other hand, if ρ is invariant, the θ-lifts are limits of discrete series. When ρ is not a discrete series, one can still expand $\operatorname{tr} \rho$ using the compatibility properties described on [Wei2, p. 4]; this is done in great detail on [Wei1, p. 93].

4.1.4 *Weak endoscopic lifts.* Following [Wei2], we say that a cuspidal automorphic representation π of $G'(\mathbb{A})$ is *endoscopic* if there exist two cuspidal automorphic representations ρ_i of $\mathrm{GL}(2)$, with the same central character ω_π, such that the L-functions satisfy

$$L(s, \pi_p, \mathrm{spin}) = L(s, \rho_{1,p}) L(s, \rho_{2,p})$$

for almost all p. Then we also say that π is a *weak* endoscopic lift of $\rho = \rho_1 \otimes \rho_2$. Moreover, let us recall what it means for π to be CAP (cuspidal associated to parabolic): π is said to be CAP with respect to a parabolic P, with Levi component M, if there exists a cuspidal automorphic representation τ of $M(\mathbb{A})$ such that π is weakly equivalent to the constituents of the induced representation of τ to $G'(\mathbb{A})$. The CAP representations for $G'(\mathbb{A})$ are described in [Pia83] and [Sou88].

THEOREM 4.4. *Let π be a weak endoscopic lift of ρ that is non-CAP. Then for all p we have*

$$\Delta_{G',H}(\rho_p, \pi_p) \neq 0.$$

Proof. This is part 3 of the main theorem on [Wei2, p. 16]. The main ingredient is a result of Kudla *et al.* [KRS92], which shows that any constituent of π restricted to $\mathrm{Sp}(4)$ is a global θ-lift from some $\mathrm{O}(X)$, since the degree-five L-function of π has a simple pole at $s = 1$. □

4.1.5 *Representations of type IIIa and θ-correspondence.* As we have shown in Theorem A, by raising the level of a suitable automorphic representation π of $G(\mathbb{A})$ we obtain a $\tilde{\pi} \equiv \pi \pmod{\lambda}$ with $\tilde{\pi}_q$ of type IIIa. This means precisely that $\tilde{\pi}_q$ is of the form

$$\chi \rtimes \sigma \mathrm{St}_{\mathrm{GL}(2)}$$

for unramified characters χ and σ of $\mathbb{Q}_q^*$ such that $\chi \neq \mathbf{1}$ and $\chi \neq |\cdot|^{\pm 2}$; they are both unitary in our case. Throughout, we use the notation of [ST93] so that the above representation is induced from the Klingen–Levi. In our case, it has trivial central character, i.e. $\chi\sigma^2 = \mathbf{1}$. We note that

$\chi^2 \neq \mathbf{1}$: indeed, $\tilde{\pi}_q$ is congruent (mod λ) to its unramified relative $\chi \rtimes \sigma \mathbf{1}_{\mathrm{GL}(2)}$, which has Satake parameters

$$\{q\alpha^{-1}, q^2\alpha^{-1}, q\alpha, q^2\alpha\}, \quad \alpha = \sigma(q),$$

up to a twist. Since $\tilde{\pi}_q \equiv \pi_q \equiv \mathbf{1}$, the above parameters are congruent to

$$\{1, q, q^2, q^3\}.$$

If $\alpha^4 = 1$, this can only happen if $q^4 \equiv 1 \pmod{\ell}$, contradicting banality. Therefore $\chi^2 \neq \mathbf{1}$, and by [ST93, Theorem 5.2(iv)] the restriction of $\tilde{\pi}_q$ to Sp(4) must remain irreducible.

Lemma 4.5. *Let $\chi \rtimes \sigma\, \mathrm{St}_{\mathrm{GL}(2)}$ be a unitary representation of $G'^{\mathrm{ad}}(\mathbb{Q}_q)$ of type IIIa, where $\chi^2 \neq \mathbf{1}$. Let X be an even-dimensional quadratic space over $\mathbb{Q}_q$ of discriminant* 1. *The representation does not occur in the θ-correspondence with* $\mathrm{GO}(X)$ *if X is anisotropic or if* $\dim X$ *is at most* 4.

Proof. By [Rob96, Lemma 4.2], it suffices to show that $\chi \rtimes \mathrm{St}_{\mathrm{SL}(2)}$ does not occur in the θ-correspondence with $\mathrm{O}(X)$. In other words, by Frobenius reciprocity, we need to show that

$$\mathrm{Hom}_{\mathrm{GL}(1)\times\mathrm{SL}(2)}(r(\omega), \chi \otimes \mathrm{St}_{\mathrm{SL}(2)}) = 0.$$

Here $r(\omega)$ is the Jacquet module for the Weil representation ω with respect to the Klingen parabolic in Sp(4). We will utilize Kudla's filtration of $r(\omega)$ as described in [Kud, Theorem 8.1]:

$$0 \to \mathrm{Ind}_P^{\mathrm{O}(X)}(\breve{\omega} \otimes \omega_{1,\breve{X}}) \to r(\omega) \to \chi_X \otimes \omega_{1,X} \to 0.$$

Here, up to a real twist, χ_X is a quadratic character. Of course, $\omega_{1,X}$ denotes the Weil representation for the pair $\mathrm{SL}(2) \times \mathrm{O}(X)$. The submodule of $r(\omega)$ is to be regarded as being trivial if X is anisotropic. Otherwise, P denotes the parabolic subgroup of $\mathrm{O}(X)$ with Levi component $\mathrm{GL}(1) \times \mathrm{O}(\breve{X})$, where $\breve{X}$ is the space in the Witt tower of X with index one less than that of X. Up to a twist, $\breve{\omega}$ is the representation of $\mathrm{GL}(1) \times \mathrm{GL}(1)$ on Schwartz functions on $\mathbb{Q}_q$ given by translation composed with multiplication. Let us first note that the following space vanishes:

$$\mathrm{Hom}_{\mathrm{GL}(1)\times\mathrm{SL}(2)}(\chi_X \otimes \omega_{1,X}, \chi \otimes \mathrm{St}_{\mathrm{SL}(2)}) = 0.$$

Otherwise, $\chi = \chi_X$; however, χ is unitary and non-quadratic. This proves the lemma when X is anisotropic. We may then assume that X is split of dimension 2 or 4. It remains to show that

$$\mathrm{Hom}_{\mathrm{GL}(1)\times\mathrm{SL}(2)}(\mathrm{Ind}_P^{\mathrm{O}(X)}(\breve{\omega} \otimes \omega_{1,\breve{X}}), \chi \otimes \mathrm{St}_{\mathrm{SL}(2)}) = 0.$$

If not, then it follows immediately from [GG05, Lemma 9.4] that $\chi \otimes \mathrm{St}_{\mathrm{SL}(2)}$ is also a quotient of the representation $\breve{\omega} \otimes \omega_{1,\breve{X}}$. Consequently, $\mathrm{St}_{\mathrm{SL}(2)}$ occurs in the θ-correspondence with $\mathrm{O}(\breve{X})$. However, it is well-known that $\mathrm{St}_{\mathrm{SL}(2)}$ does not come from split O(2); see the example on [Kud, p. 86]. □

Corollary 4.6. *Let π be a cuspidal automorphic representation of $G'^{\mathrm{ad}}(\mathbb{A})$ having a local component of type IIIa of the form $\chi \rtimes \sigma \mathit{St}_{\mathrm{GL}(2)}$ where $\chi^2 \neq \mathbf{1}$. Then π is neither CAP nor endoscopic.*

Proof. Suppose π is CAP with respect to the Siegel parabolic P_β or B. Note that PGSp(4) is the same as split SO(5); then, by [Pia83], it comes from $\widetilde{\mathrm{SL}}(2)$ via global θ-lifting. Locally, one can compute these θ-lifts and check that they are all non-generic (we will have more to say about this in the next section). However, type IIIa representations are generic. Now suppose that π

is CAP with respect to the Klingen parabolic P_α. By [Sou88], there exists a two-dimensional anisotropic quadratic space X over $\mathbb{Q}$ such that π is a global θ-lift from $\mathrm{GO}(X)$. However, by Lemma 4.5, type-IIIa representations do not occur in the θ-correspondence with any two-dimensional quadratic space. Finally, suppose π is a weak endoscopic lift of ρ. By Theorem 4.4, the local component $\pi_q = \chi \rtimes \sigma \mathrm{St}_{\mathrm{GL}(2)}$ occurs in the expansion of tr ρ_q. If ρ_q is a discrete series, this is impossible by Theorem 4.3 and Lemma 4.5; otherwise, its character expansion is given by a single representation (see [Wei1, p. 94]). This representation is irreducibly induced from P_β or B, and thus it cannot be of type IIIa. □

4.2 Stability

4.2.1 *Stabilization of the trace formula.* The trace formula for G' is an equality between two expansions of a very complicated invariant distribution $I^{G'}$ on $G'(\mathbb{A})$. One expansion is in terms of geometric data such as conjugacy classes, Tamagawa numbers, and (weighted) orbital integrals. The other expansion is in terms of spectral data such as automorphic representations, multiplicities, and (weighted) characters. For our purpose, we are only interested in the terms occurring discretely in the trace formula. Their sum defines an invariant distribution denoted by $I^{G'}_{\mathrm{disc}}$. The main contribution comes from the trace on the discrete spectrum, but there are also terms coming from what Arthur refers to as surviving remnants of Eisenstein series. The distribution has an expansion of the form

$$I^{G'}_{\mathrm{disc}}(f') = \sum_{\Pi} a^{G'}_{\mathrm{disc}}(\Pi) \mathrm{tr}\, \Pi(f')$$

for a smooth function f' on $G'(\mathbb{A})$. Here $a^{G'}_{\mathrm{disc}}(\Pi)$ is a complex number attached to the discrete automorphic representation Π. If Π is cuspidal but not CAP, the number $a^{G'}_{\mathrm{disc}}(\Pi)$ is simply the multiplicity of Π. The distribution $I^{G'}_{\mathrm{disc}}$ is unstable (recall that a distribution is said to be stable if it is supported on the stable orbital integrals). However, by the work of Arthur announced in [Art98],

$$S^{G'}_{\mathrm{disc}}(f') = I^{G'}_{\mathrm{disc}}(f') - \tfrac{1}{4} I^{H}_{\mathrm{disc}}(f'^{H})$$

does define a stable distribution. Here, if $f' = \otimes f'_p$ is a pure tensor, we may take the matching function to be $f'^H = \otimes f'^H_p$. Now we turn our attention to the trace formula for G. Since G is anisotropic modulo its center, the trace formula takes its simplest form. All terms occur discretely, that is,

$$I^{G}_{\mathrm{disc}}(f) = \sum_{\pi} a^{G}_{\mathrm{disc}}(\pi) \mathrm{tr}\, \pi(f)$$

for a smooth function f on $G(\mathbb{A})$. Here $a^{G}_{\mathrm{disc}}(\pi)$ is always the multiplicity of π. Again, this distribution is unstable, but it can be rewritten in terms of stable distributions on the endsocopic groups:

$$I^{G}_{\mathrm{disc}}(f) = S^{G'}_{\mathrm{disc}}(f^{G'}) + \tfrac{1}{4} I^{H}_{\mathrm{disc}}(f^{H}).$$

This was first proved by Kottwitz and Langlands, but it is also a very special case of the aforementioned work of Arthur. If $f = \otimes f_p$ is a tensor product, we may take $f^{G'} = \otimes f^{G'}_p$ as before.

4.2.2 *A semilocal spectral identity.* As already observed, we have global transfer. For example, if f is a function on $G(\mathbb{A})$, there is a function f^H on $H(\mathbb{A})$ with matching orbital integrals. There is also a global character identity,

$$\mathrm{tr}\,\rho(f^H) = \sum_{\pi} \Delta_{G,H}(\rho,\pi)\mathrm{tr}\,\pi(f)$$

for any $f \in C_c^\infty(G(\mathbb{A}))$, where ρ is an irreducible admissible representation of $H(\mathbb{A})$. In the sum, π runs over irreducible admissible representations of $G(\mathbb{A})$, and $\Delta_{G,H}(\rho,\pi)$ is the product of the local transfer factors $\Delta_{G,H}(\rho_p,\pi_p)$. If we insert this expansion into the stable trace formula, we see that

$$\sum_{\pi}\left\{a_{\mathrm{disc}}^{G}(\pi) - \frac{1}{4}\sum_{\rho} a_{\mathrm{disc}}^{H}(\rho)\Delta_{G,H}(\rho,\pi)\right\}\mathrm{tr}\,\pi(f)$$

equals

$$\sum_{\Pi}\left\{a_{\mathrm{disc}}^{G\prime}(\Pi) - \frac{1}{4}\sum_{\rho} a_{\mathrm{disc}}^{H}(\rho)\Delta_{G',H}(\rho,\Pi)\right\}\mathrm{tr}\,\Pi(f')$$

for any pair of matching functions f and f'. We want to refine this identity. The point is that G is split over $\mathbb{Q}_p$ for all $p \notin S$. Thus, if we fix an irreducible representation τ^S of the group $G(\mathbb{A}^S)$,

$$\sum_{\pi_S}\left\{a_{\mathrm{disc}}^{G}(\pi_S\otimes\tau^S) - \frac{1}{4}\sum_{\rho} a_{\mathrm{disc}}^{H}(\rho)\Delta_{G,H}(\rho,\pi_S\otimes\tau^S)\right\}\mathrm{tr}\,\pi_S(f_S)$$

equals

$$\sum_{\Pi_S}\left\{a_{\mathrm{disc}}^{G\prime}(\Pi_S\otimes\tau^S) - \frac{1}{4}\sum_{\rho} a_{\mathrm{disc}}^{H}(\rho)\Delta_{G',H}(\rho,\Pi_S\otimes\tau^S)\right\}\mathrm{tr}\,\Pi_S(f_S')$$

for any pair of matching functions f_S and f_S', by linear independence of characters for $G(\mathbb{A}^S)$. From now on, we assume that τ^S comes from an automorphic representation τ of $G(\mathbb{A})$ such that

$$\Delta_{G,H}(\rho_p,\tau_p) = 0 \quad \text{for some } p \notin S,$$

for every discrete automorphic representation ρ of $H(\mathbb{A})$. This is true, for example, if τ^S has a local component of type IIIa as above. Under this assumption, the above identity simplifies immensely to

$$\sum_{\pi_S} a_{\mathrm{disc}}^{G}(\pi_S\otimes\tau^S)\mathrm{tr}\,\pi_S(f_S) = \sum_{\Pi_S} a_{\mathrm{disc}}^{G\prime}(\Pi_S\otimes\tau^S)\mathrm{tr}\,\Pi_S(f_S')$$

for any pair of matching functions f_S and f_S'. Let us mention that if the above hypothesis on τ^S does not hold, then there exists a ρ such that $\Delta_{G,H}(\rho_p,\tau_p)$ is non-zero for all $p \notin S$. We may then construct a weak transfer of τ to $G'(\mathbb{A})$ by looking at the global θ-lift of ρ as in [Wei2].

4.2.3 *Incorporating Shelstad's results at infinity.* For now, let us fix a pair of matching functions f_r and f_r' at r, and consider

$$T = \sum_{\Pi_S} a_{\mathrm{disc}}^{G\prime}(\Pi_S\otimes\tau^S)\mathrm{tr}\,\Pi_r(f_r')\mathrm{tr}\,\Pi_\infty.$$

This is a distribution on $G'(\mathbb{R})$. From our previous considerations, this is clearly stable. Recall that by the Langlands classification, the irreducible admissible representations of $G'(\mathbb{R})$ are partitioned into finite L-packets Π_μ parameterized by admissible homomorphisms $\mu : W_{\mathbb{R}} \to {}^LG$.

Then, by [She82], T has an expansion

$$T = \sum_{\mu} c_\mu \mathrm{tr}\, \Pi_\mu, \quad \mathrm{tr}\, \Pi_\mu = \sum_{\Pi_\infty \in \Pi_\mu} \mathrm{tr}\, \Pi_\infty,$$

where μ varies over the *tempered* parameters, i.e. all μ such that the projection of $\mu(W_{\mathbb{R}})$ onto the neutral component of ${}^L G$ is bounded. Indeed, we assume that τ^S is part of a non-CAP representation, so it is tempered by [Wei05]. From an argument in [BR94] using the congruence relation, it follows that Π_∞ must be tempered for $\Pi_S \otimes \tau^S$ to be cuspidal. For every tempered μ, the coefficient c_μ is given by

$$c_\mu = \sum_{\Pi_r} a^{G'}_{\mathrm{disc}}(\Pi_\infty \otimes \Pi_r \otimes \tau^S) \mathrm{tr}\, \Pi_r(f'_r)$$

for any choice $\Pi_\infty \in \Pi_\mu$. Now, since $G(\mathbb{R})$ is compact, its L-packets are singletons $\{\pi_\mu\}$. The finite-dimensional irreducible representations π_μ are parameterized by discrete L-parameters μ (i.e. any μ which does not map into a Levi subgroup). In this case, the L-packet Π_μ for $G'(\mathbb{R})$ consists of two classes of discrete series representations $\{\Pi^H_\mu, \Pi^W_\mu\}$ with the same central and infinitesimal characters as π_μ. The representation Π^H_μ is a holomorphic discrete series, whereas Π^W_μ is generic. We will now invoke the following character identity over $\mathbb{R}$ proved by Shelstad in [She79]:

$$\mathrm{tr}\, \Pi_\mu(f'_\infty) = \begin{cases} \mathrm{tr}\, \pi_\mu(f_\infty) & \text{if } \mu \text{ is discrete,} \\ 0 & \text{otherwise,} \end{cases}$$

for matching functions f_∞ and f'_∞. Inserting this into the trace formula derived in the last section, we obtain

$$c_\mu = \sum_{\pi_r} a^{G}_{\mathrm{disc}}(\pi_\mu \otimes \pi_r \otimes \tau^S) \mathrm{tr}\, \pi_r(f_r)$$

for any discrete μ. To elaborate on this, we compute $T(f'_\infty)$ in two ways and then use linear independence of characters for $G(\mathbb{R})$. Comparing this with the above, we obtain our *key* identity

$$\sum_{\pi_r} a^{G}_{\mathrm{disc}}(\pi_\mu \otimes \pi_r \otimes \tau^S) \mathrm{tr}\, \pi_r(f_r) = \sum_{\Pi_r} a^{G'}_{\mathrm{disc}}(\Pi_\infty \otimes \Pi_r \otimes \tau^S) \mathrm{tr}\, \Pi_r(f'_r),$$

which is valid for any discrete μ, any $\Pi_\infty \in \Pi_\mu$, and any matching pair f_r and f'_r at r.

4.3 Proof of Theorem C

Let τ be an automorphic representation of $G(\mathbb{A})$ having a local component of type IIIa outside S. Suppose $\tau_\infty = \pi_\mu$. Then, by linear independence of characters for $G(\mathbb{Q}_r)$, there exists a function f_r such that the left-hand side of the key identity above is non-zero. Let f'_r be any matching function. Then the right-hand side is non-zero, and there exists a Π_r with $\mathrm{tr}\, \Pi_r(f'_r) \neq 0$ such that $\Pi_\infty \otimes \Pi_r \otimes \tau^S$ is a discrete automorphic representation of $G'(\mathbb{A})$; call it Π. It has a tempered component (namely the one of type IIIa), so Π must in fact be cuspidal. This is a standard argument using the fact, proved by Langlands [Lan76], that residual representations arise from residues of Eisenstein series for non-unitary parameters. The same argument is used in a paper of Labesse and Muller [LM04, p. 6]. As we have shown, since Π has a component of type IIIa, it is neither CAP nor endoscopic. Finally, we note that our argument can be extended to allow central characters.

To conclude, we refine our argument to gain information at the prime r. Let τ be as above, but insist that $\omega_\tau = \mathbf{1}$ and that τ_r is para-spherical. This means $\tau_r^{K_r} \neq 0$ for a paramodular group K_r

in $G(\mathbb{Q}_r)$. If we take $f_r = e_{K_r}$, the left-hand side of the key identity is positive. By Theorem B, we may then take $f'_r = e_{\eta K'_r}$. Hence there exists a cuspidal automorphic representation Π of $G'^{\mathrm{ad}}(\mathbb{A})$ with $\Pi_\infty \in \Pi_\mu$ and $\Pi^S = \tau^S$ such that the trace of η on $\Pi_r^{K'_r}$ is positive. In particular, Π_r is para-spherical. We claim that Π_r is also ramified. Suppose not; then since Π is not CAP, Π_r is tempered, by [Wei05, Theorem I]. Thus Π_r must be a full unramified principal series. However, the Atkin–Lehner operator on $\Pi_r^{K'_r}$ would then be traceless by [Sch05, Table 3], and this is a contradiction. □

Remark 1. The aforementioned table yields that Π_r must be of type IIa, Vb, Vc or VIc. We suspect that Π_r is necessarily tempered. If this is true, we deduce that Π_r is of type IIa of the form $\chi\mathrm{St} \rtimes \sigma$ (induced from the Siegel parabolic) where $\chi\sigma$ is the non-trivial unramified quadratic character of $\mathbb{Q}_r^*$.

5. Saito–Kurokawa forms

5.1 Modular forms and root numbers

Let $f \in S_4(\Gamma_0(N))$ be a newform of square-free level N, and consider its L-function given by the usual Euler product: for a_n the nth Hecke eigenvalue of f and for $\mathrm{Re}(s)$ sufficiently large,

$$L(s,f) = \sum_{n=1}^{\infty} a_n n^{-s} = \prod_{p|N}(1-a_p p^{-s})^{-1} \cdot \prod_{p\nmid N}(1-a_p p^{-s}+p^{3-2s})^{-1}.$$

This function has analytic continuation to the s-plane and satisfies the functional equation

$$\Lambda(s,f) = (2\pi)^{-s}\Gamma(s)L(s,f) = \epsilon_f N^{2-s}\Lambda(4-s,f),$$

where the root number $\epsilon_f \in \{\pm 1\}$ is given by the parity of the order of vanishing of $L(s,f)$ at the point $s=2$. Now, we wish to work in the context of automorphic representations. By an elementary construction, one associates to f a cuspidal automorphic representation τ of $\mathrm{PGL}(2,\mathbb{A})$. Specifically, one first pulls back f to a function on $\mathrm{GL}(2,\mathbb{R})^+$, the neutral component; then, by strong approximation, one views it as a function on $\mathrm{GL}(2,\mathbb{A})$, and τ is the representation it generates. The representation τ is uniquely determined by the following properties: τ_p is unramified for $p \nmid N$, and its Satake parameters $\{\alpha_p, \alpha_p^{-1}\}$ satisfy

$$a_p = p^{3/2}(\alpha_p + \alpha_p^{-1}).$$

Moreover, up to an appropriate twist, τ_∞ is the (holomorphic) discrete series representation of $\mathrm{GL}(2,\mathbb{R})$ with the same central and infinitesimal characters as $\mathrm{Sym}^2(\mathbb{C}^2)$. For p dividing N, the component τ_p is in fact an unramified quadratic twist of $\mathrm{St}_{\mathrm{GL}(2)}$, since N is assumed to be square-free. We note that the Jacquet–Langlands L-function $L(s,\tau)$ is simply $\Lambda(s+3/2,f)$. In addition, $\epsilon_f N^{1/2-s}$ equals the exponential function $\epsilon(s,\tau)$ in its functional equation. Thus

$$\epsilon(1/2,\tau) = \epsilon_f.$$

5.2 The Saito–Kurokawa lifting

Let f be as above, but now assume that $\epsilon_f = -1$. We then lift τ to $\mathrm{PGSp}(4)$.

Proposition 5.1. *There exists a cuspidal automorphic representation Π of $\mathrm{PGSp}(4)$, with Π_∞ being the cohomological holomorphic discrete series representation, such that for all primes p we* *have*

$$\Pi_p \simeq L(\nu^{1/2}\tau_p \rtimes \nu^{-1/2}).$$

Here ν denotes the normalized absolute value, and $L(-)$ is the unique irreducible quotient. In particular, Π_p is of type IIb for $p \nmid N$. On the other hand, for p dividing N, Π_p is of type VIc or Vb according to whether τ_p is $\mathrm{St}_{\mathrm{GL}(2)}$ or its non-trivial unramified quadratic twist $\xi_0\mathrm{St}_{\mathrm{GL}(2)}$.

Proof. Let $\widetilde{\mathrm{SL}}(2)$ denote the twofold metaplectic covering of $\mathrm{SL}(2)$. Throughout, we also fix a non-trivial additive unitary character $\psi = \otimes\psi_p$. Each τ_p is infinite-dimensional, so it determines a local Waldspurger packet A_{τ_p} of irreducible unitary representations of $\widetilde{\mathrm{SL}}(2)$ over $\mathbb{Q}_p$. This packet is a singleton $\{\sigma_{\tau_p}^+\}$ when τ_p is a principal series. Otherwise, when τ_p is a discrete series,

$$A_{\tau_p} = \{\sigma_{\tau_p}^+, \sigma_{\tau_p}^-\}.$$

Here $\sigma_{\tau_p}^+$ is ψ_p-generic, whereas $\sigma_{\tau_p}^-$ is not. Recall that $\mathrm{PGL}(2)$ is the same as split $\mathrm{SO}(3)$, and its inner form PD^* is anisotropic $\mathrm{SO}(3)$. Then $\sigma_{\tau_p}^+$ can be described as the θ-lift of τ_p. In the discrete-series case, $\sigma_{\tau_p}^-$ is the θ-lift of the Jacquet–Langlands transfer $\breve{\tau}_p$. Consider the tensor product

$$A_\tau = \otimes A_{\tau_p} = \{\sigma = \otimes\sigma_{\tau_p}^{\epsilon_p} \text{ with } \epsilon_p = \pm \text{ and } \epsilon_p = + \text{ for almost all } p\}.$$

This is the global Waldspurger packet determined by τ. It is a finite set of irreducible unitary representations of $\widetilde{\mathrm{SL}}(2)$ over $\mathbb{A}$. They are not all automorphic; the signs have to be compatible. Indeed, by a famous result of Waldspurger,

$$\sigma = \otimes\sigma_{\tau_p}^{\epsilon_p} \text{ is automorphic } \Leftrightarrow \epsilon(1/2, \tau) = \prod \epsilon_p.$$

For example, in our case, $\sigma = \sigma_{\tau_\infty}^- \otimes_{p<\infty} \sigma_{\tau_p}^+$ is automorphic since $\epsilon_f = -1$. Now we think of $\mathrm{PGSp}(4)$ as split $\mathrm{SO}(5)$ and look at the global θ-series lifting $\theta(\sigma)$. This is non-zero. Indeed, we are in the stable range. We claim that $\theta(\sigma)$ is contained in the space of cusp forms. Otherwise, by the theory of towers due to Rallis, σ would have a cuspidal lift to $\mathrm{PGL}(2)$. However, a result of Waldspurger then implies that σ is generic; on the other hand, σ_∞ is non-generic. From a short argument (see, for example, [Gan08, Proposition 2.12] and its proof) it then follows that we have local–global compatibility, that is,

$$\Pi = \theta(\sigma) = \theta(\sigma_{\tau_\infty}^-) \otimes_{p<\infty} \theta(\sigma_{\tau_p}^+).$$

In particular, $\theta(\sigma)$ is irreducible. We should mention that local Howe duality is known in this special situation. The case $p = 2$ can be checked by hand. It remains to compute the local lifts above. Using [Kud, Proposition 4.1], it is not hard to show that $\theta(\sigma_{\tau_p}^+)$ is the Langlands quotient given in our proposition. Furthermore, by [Li90], $\theta(\sigma_{\tau_\infty}^-)$ is the holomorphic discrete series with minimal K-type $(3, 3)$. □

It follows immediately that Π is of Saito–Kurokawa type (that is, CAP with respect to the Siegel parabolic). Moreover, it is non-tempered and para-spherical at all finite primes.

5.3 Transferring to an inner form

We now assume that $\psi = \otimes\psi_p$ has trivial conductor. Then, as is well-known,

$$\epsilon(1/2, \mathrm{St}_{\mathrm{GL}(2)}, \psi_p) = -1, \quad \epsilon(1/2, \xi_0\mathrm{St}_{\mathrm{GL}(2)}, \psi_p) = +1.$$

We have $N > 1$, so we pick a prime r such that $\tau_r = \mathrm{St}_{\mathrm{GL}(2)}$. Then let D be the division quaternion algebra over $\mathbb{Q}$ with ramification locus $S = \{\infty, r\}$, and let G be the unitary similitude group of D^2. The reduced norm of D_r maps onto $\mathbb{Q}_r$, so all hermitian forms on D_r^2 are equivalent. For example, we may take

$$G(\mathbb{Q}_r) = \left\{ x \in \mathrm{GL}(2, D_r) : x^* \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} x = c(x) \begin{pmatrix} & 1 \\ 1 & \end{pmatrix},\ c(x) \in \mathbb{Q}_r^* \right\}.$$

Consider the isotropic subspace $D_r \oplus 0$. Its stabilizer is the minimal parabolic

$$P = \left\{ \begin{pmatrix} a & \\ & c\bar{a}^{-1} \end{pmatrix} : a \in D_r^* \text{ and } c \in \mathbb{Q}_r^* \right\} \ltimes \left\{ \begin{pmatrix} 1 & b \\ & 1 \end{pmatrix} : b + \bar{b} = 0 \right\}.$$

It has Levi component $D_r^* \times \mathbb{Q}_r^*$ and abelian unipotent radical. Furthermore,

$$\delta_P : \begin{pmatrix} a & \\ & c\bar{a}^{-1} \end{pmatrix} \mapsto |N_{D_r/\mathbb{Q}_r}(a)|^3 \cdot |c|^{-3},$$

as can be shown by a standard calculation. We now transfer Π to G using θ-correspondence.

Proposition 5.2. *There exists an automorphic representation π of $G(\mathbb{A})$, with ω_π and π_∞ being trivial, which agrees with Π outside of S and is such that the local component at the ramified prime r is*

$$\pi_r \simeq \nu^{1/2}\mathbf{1}_{D^*} \rtimes \nu^{-1/2}.$$

Proof. We use the notation from the proof of the previous proposition. Since τ_r is a discrete series representation, the global Waldspurger packet A_τ contains another automorphic member, namely

$$\breve{\sigma} = \sigma^+_{\tau_\infty} \otimes \sigma^-_{\tau_r} \otimes_{p \neq r} \sigma^+_{\tau_p}.$$

Now we realize the adjoint group G^{ad} as a certain anisotropic SO(5) and look at the global θ-lift to this group $\theta(\breve{\sigma})$. We are no longer in the stable range, so to make sure that this is non-vanishing, we appeal to the Rallis inner-product formula. The case we need is reviewed on [Gan08, p. 9]. Our quadratic space has dimension 5, so all special L-values in the inner product formula are non-zero. Consequently, $\theta(\breve{\sigma}) \neq 0$ if and only if all the local lifts $\theta(\breve{\sigma}_p)$ are non-vanishing. However,

$$\theta(\sigma^+_{\tau_\infty}) = \mathbf{1}, \quad \theta(\sigma^-_{\tau_r}) = L(\nu^{1/2}\mathbf{1}_{D^*} \rtimes \nu^{-1/2}).$$

The first identity is a consequence of [Kud, Theorem 5.1], and the second is easily derived from [Kud, Proposition 4.1]. As before, we have local–global compatibility, and we take $\pi = \theta(\breve{\sigma})$. It remains to show that the unramified principal series $\nu^{1/2}\mathbf{1}_{D^*} \rtimes \nu^{-1/2}$ is irreducible. This is an easy exercise using the expression for δ_P and the results of Kato reviewed on [Car79, p. 144]. □

We note in passing that the unramified principal series $\nu^{1/2}\xi_0\mathbf{1}_{D^*} \rtimes \nu^{-1/2}$ is reducible. Therefore it is crucial that we pick r such that τ_r is the actual Steinberg representation $\mathrm{St}_{\mathrm{GL}(2)}$ and not its twist $\xi_0\mathrm{St}_{\mathrm{GL}(2)}$. The following lemma allows us to apply Theorem A to raise the level of π.

Lemma 5.3. *The representation π occurs with multiplicity one.*

Proof. We first recall that for $\widetilde{\mathrm{SL}}(2, \mathbb{A})$, Waldspurger proved multiplicity one. Essentially, this follows from the multiplicity-one theorem for $\mathrm{PGL}(2, \mathbb{A})$, using the θ-correspondence. We can therefore identify the abstract representation $\breve{\sigma}$ with a space of cusp forms on the metaplectic group. By a formal argument like the one on [Gan08, p. 8], the θ-correspondence preserves multiplicity. Thus, the representation $\pi = \theta(\breve{\sigma})$ occurs with multiplicity one in the spectrum of G. □

Gan proves a more general result in [Gan08]. Analogous to work of Piatetski-Shapiro and Sayag in the isotropic case, Gan characterizes (certain) CAP representations of an anisotropic inner form of GSp(4) as θ-lifts from the metaplectic group. As a corollary, he deduces that all these CAP representations have multiplicity one. We will use this characterization later.

Let us end this section with a few words about the Bruhat–Tits theory of $G(\mathbb{Q}_r)$. We denote by $\mathcal{O}_{D_r}$ the maximal compact subring of D_r, and we let $\mathfrak{p}_{D_r}$ be its (bilateral) maximal ideal. We choose a uniformizing parameter ϖ_{D_r}. The order $\mathcal{O}_{D_r}$ defines an integral model for $G_{/\mathbb{Q}_r}$, and we introduce the subgroup

$$K = G(\mathbb{Z}_r) = \left\{ x \in \mathrm{GL}(2, \mathcal{O}_{D_r}) : x^* \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} x = c(x) \begin{pmatrix} & 1 \\ 1 & \end{pmatrix},\ c(x) \in \mathbb{Z}_r^* \right\}.$$

This is the special maximal compact subgroup of $G(\mathbb{Q}_r)$, which becomes the Siegel parahoric over the unramified quadratic extension of $\mathbb{Q}_r$. Inside K, there is the Iwahori subgroup I consisting of matrices which are upper triangular modulo $\mathfrak{p}_{D_r}$. We then let $\tilde{K}$ be the subgroup of G generated by I and

$$\begin{pmatrix} & \varpi_{D_r}^{-1} \\ \varpi_{D_r} & \end{pmatrix}.$$

This $\tilde{K}$ is the paramodular maximal compact subgroup of G. Both K and $\tilde{K}$ are special, so they fit into Iwasawa decompositions of G relative to P. Consequently, the representation π_r is both K-spherical and $\tilde{K}$-spherical.

5.4 Galois representations

Let L_f be the number field generated by the Hecke eigenvalues of f. A classical construction due to Deligne [Del71], generalizing Eichler–Shimura theory, provides a compatible system of continuous irreducible Galois representations

$$\rho_{f,\lambda} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(2, L_{f,\lambda}),$$

indexed by the places $\lambda | \ell$ of L_f, such that $\rho_{f,\lambda}$ is unramified at $p \nmid N\ell$ and

$$L_p(s, f) = \det(1 - \rho_{f,\lambda}(\mathrm{Frob}_p)p^{-s})^{-1}$$

for such p. Here Frob_p denotes a geometric Frobenius. This result has been generalized to GSp(4) by Laumon [Lau97] and Weissauer [Wei05]. Specifically, suppose Π is a cuspidal automorphic representation of GSp(4) with Π_∞ being a cohomological discrete series; then there exists a number field L_Π and a compatible system of continuous semisimple four-dimensional Galois representations

$$\rho_{\Pi,\lambda} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(4, L_{\Pi,\lambda}),$$

indexed by the finite places $\lambda | \ell$ of L_Π, such that $\rho_{\Pi,\lambda}$ is unramified at $p \neq \ell$ outside the ramification locus of Π. Moreover, for such primes p there is the following relation with the spinor L-factor:

$$L(s - w/2, \Pi_p, \mathrm{spin}) = \det(1 - \rho_{\Pi,\lambda}(\mathrm{Frob}_p)p^{-s})^{-1}.$$

Here $w = k_1 + k_2 - 3$, where (k_2, k_2) is the weight of Π_∞. We note that when Π is not CAP, the representation $\rho_{\Pi,\lambda}$ can be shown (see [Wei05]) to be pure of weight w. This means that the eigenvalues of Frob_p have absolute value $p^{w/2}$. When Π is CAP or endsocopic, $\rho_{\Pi,\lambda}$ is reducible and essentially given by the above construction of Deligne. For example, for our Saito–Kurokawa form, it is given by the following lemma.

Lemma 5.4. *Let* Π *be the Saito–Kurokawa lift of* f *as in Proposition 5.1.*
Then

$$\rho_{\Pi,\lambda} \simeq \rho_{f,\lambda} \oplus \omega_\ell^{-1} \oplus \omega_\ell^{-2}$$

for all $\lambda|\ell$*, where* ω_ℓ *denotes the* ℓ*-adic cyclotomic character.*

Proof. Suppose $p \nmid N\ell$ and that τ_p is induced from the unramified character χ:

$$\Pi_p \simeq L(\nu^{1/2}\chi \times \nu^{1/2}\chi^{-1} \rtimes \nu^{-1/2}) \simeq \chi \mathbf{1}_{\mathrm{GL}(2)} \rtimes \chi^{-1}.$$

This allows us to calculate the Satake parameters, and it follows that

$$L(s, \Pi_p, \mathrm{spin}) = L(s, \tau_p)\zeta_p(s-1/2)\zeta_p(s+1/2).$$

In our case, $k_1 = k_2 = 3$, so $w = 3$. Now use the fact that $\omega_\ell(\mathrm{Frob}_p) = p^{-1}$ for $p \neq \ell$. □

In contrast, when Π is *neither* CAP *nor* endoscopic, the Galois representation $\rho_{\Pi,\lambda}$ is expected to be irreducible. For large ℓ there is the following precise result in this direction.

Theorem 5.5. *Let* Π *be a cuspidal automorphic representation of* GSp(4) *with* Π_∞ *cohomological. Assume that* $\ell > 2w+1$*. Suppose* $\rho_{\Pi,\lambda}$ *is reducible and that all of its two-dimensional constituents are odd. Then* Π *is CAP or endoscopic.*

Proof. See [SU06, Theorem 3.2.1]. It relies on work of Ramakrishnan [Ram08]. □

5.5 Proof of Theorem D

Let $K = \prod K_p$, where K_p is hyperspecial for $p \nmid N$ and paramodular for $p|N$. We want to apply Theorem A to the automorphic representation π from Proposition 5.2. However, we cannot prove that K is a good small subgroup (in the sense that π_f^K determines π_f). Indeed, the paramodular groups do not have Iwahori factorizations with respect to any parabolic. There is a way to circumvent this problem: all we need is that the module π_f^K has multiplicity one in $\mathcal{A}_K$. To see this, suppose that

$$\pi_f'^K \simeq \pi_f^K$$

for an automorphic representation π' with $\omega_{\pi'}$ and π'_∞ trivial. We wish to show that $\pi' \simeq \pi$. Our claim then follows from Lemma 5.3. Clearly, $\pi'_p \simeq \pi_p$ for $p \nmid N$. Thus π' is weakly equivalent to the CAP representation Π of $G'(\mathbb{A})$. Then, by [Gan08, Theorem 7.1], we have $\pi' = \theta(\sigma')$ for some automorphic representation σ' in the Waldspurger packet $A_{\tau'}$ for some τ'. Here τ' must be weakly equivalent to τ. Hence $\tau' \simeq \tau$ by strong multiplicity-one for GL(2). Now consider a prime $p|N$; then $\pi'_p = \theta(\sigma^\pm_{\tau_p})$. First, look at the case where $p \neq r$. Here $\theta(\sigma^-_{\tau_p})$ is supercuspidal or of type VIb; see, for example, [Gan08, Proposition 5.5]. Both are para-ramified, so $\pi'_p = \pi_p$. Finally, let $p = r$. Since τ_r is the Steinberg representation, $\theta(\sigma^+_{\tau_r}) = 0$ by [Gan08, Proposition 6.5]. Therefore $\pi'_r = \pi_r$.

Now we apply Theorem A to π. Let $\lambda|\ell$ be a finite place of $\bar{\mathbb{Q}}$, with ℓ not dividing the discriminant of $H_{K,\mathbb{Z}}$, such that $\bar\rho_{f,\lambda}$ is irreducible. Then π is non-abelian modulo λ (otherwise Π would be congruent to an automorphic character, and its Galois representation would be a sum of characters modulo λ, contradicting Lemma 5.4).

Now suppose that $q \nmid N\ell$ is a prime number satisfying:

- $q^i \neq 1 \pmod{\ell}$ for $i = 1, \ldots, 4$;
- $\bar\rho_{f,\lambda}(\mathrm{Frob}_q)$ has a fixed vector.

The Satake parameters of π_q are $\{\alpha_q, q^{-1/2}, q^{1/2}, \alpha_q^{-1}\}$. Since $\bar{\rho}_{f,\lambda}(\mathrm{Frob}_q)$ has eigenvalues $\{1, q^3\}$, the level-raising condition in Theorem A is satisfied. As a result, we find an automorphic representation $\tilde{\pi} \equiv \pi \pmod{\lambda}$ of $G(\mathbb{A})$, with $\omega_{\tilde{\pi}}$ and $\tilde{\pi}_\infty$ trivial, such that $\tilde{\pi}_f^{Kq} \neq 0$ and $\tilde{\pi}_q$ is of type IIIa. Next, we apply Theorem C to $\tilde{\pi}$. As we have seen earlier, $\tilde{\pi}_q$ must have the form $\chi \rtimes \sigma\mathrm{St}_{\mathrm{GL}(2)}$ with $\chi^2 \neq \mathbf{1}$. Pick a cohomological discrete series representation $\Pi_{\mathbf{1}}$ of $G'(\mathbb{R})$, holomorphic or generic of weight $(3,3)$. Then we find a cuspidal automorphic representation $\tilde{\Pi}$ of $G'^{\mathrm{ad}}(\mathbb{A})$, with $\tilde{\Pi}_\infty = \Pi_{\mathbf{1}}$, such that $\tilde{\Pi}_p = \tilde{\pi}_p$ for $p \neq r$. Moreover, $\tilde{\Pi}_r$ is para-spherical since $\tilde{\pi}_r$ is. Thus $\tilde{\Pi}_p$ is para-spherical for all $p|N$, unramified and tempered (see [Wei05]) for $p \nmid Nq$, and of type IIIa for $p = q$. Obviously, $\tilde{\Pi}_p \equiv \Pi_p \pmod{\lambda}$ for almost all p. Therefore the Galois representations $\rho_{\tilde{\Pi},\lambda}$ and $\rho_{\Pi,\lambda}$ have the same semisimplifications modulo λ. In other words,

$$\rho_{\tilde{\Pi},\lambda} \simeq \bar{\rho}_{f,\lambda} \oplus \bar{\omega}_\ell^{-1} \oplus \bar{\omega}_\ell^{-2}$$

by Lemma 5.4.

It remains to show that $\rho_{\tilde{\Pi},\lambda}$ is irreducible. Suppose it is reducible; then it must be a sum $\varrho \oplus \varrho'$ of a pair of two-dimensional representations. Interchanging the two, we may assume that

$$\bar{\varrho} \simeq \bar{\rho}_{f,\lambda}, \quad \bar{\varrho}' \simeq \bar{\omega}_\ell^{-1} \oplus \bar{\omega}_\ell^{-2}.$$

Then, clearly, ϱ and ϱ' are both odd. Theorem 5.5 applies for $\ell > 7$. Hence $\tilde{\Pi}$ is CAP or endoscopic, contradicting Theorem C. This proves irreducibility of $\rho_{\tilde{\Pi},\lambda}$ and finishes the proof of Theorem D.

5.6 Existence of good primes

To apply Theorem D, we need to know the existence of primes q where we can raise the level. Assume that $\ell > 13$; then choose $g \in \mathbb{Z}$ prime to ℓ, which is a generator for $\mathbb{F}_\ell^*$ modulo ℓ. Thus

$$g^i \neq 1 \pmod{\ell}$$

for $i = 1, \dots, 12$. If f is not CM, in a suitable basis the image of $\rho_{f,\lambda}$ contains

$$\{x \in \mathrm{GL}(2, \mathbb{Z}_\ell) : \det x \in (\mathbb{Z}_\ell^*)^3\},$$

by [Rib85, Theorem 3.1]. In particular, the diagonal matrix with entries $\{1, g^3\}$ lies in the image of $\bar{\rho}_{f,\lambda}$. Then, by the Chebotarev density theorem, there exists a positive density of primes $q \nmid N\ell$ such that $\bar{\rho}_{f,\lambda}(\mathrm{Frob}_q)$ has eigenvalues $\{1, g^3\}$. The determinant is q^3, so we must have $g = \zeta q$ for some $\zeta \in \mathbb{F}_\ell$ with $\zeta^3 = 1$. If $q^i \equiv 1 \pmod{\ell}$ for some $i = 1, \dots, 4$, we deduce that $g^{3i} \equiv 1 \pmod{\ell}$.

6. The Bloch–Kato conjecture

6.1 An application of Theorem D

We continue to let $f \in S_4(\Gamma_0(N))$ be a newform of square-free level N, not of CM type, having root number $\epsilon_f = -1$. This sign condition implies that the L-function of f vanishes at the critical center $s = 2$ (under the classical normalization of the functional equation $s \mapsto 4 - s$). In this situation, a conjecture of Bloch and Kato [BK90, p. 376] predicts that an associated Selmer group is positive-dimensional. This expectation was proved in [SU06] for when ℓ is ordinary for f (i.e. $a_\ell(f)$ is an ℓ-adic unit), and our object in this section is to make progress on the conjecture when ℓ is supersingular. Let $\lambda|\ell$ be a finite place of $\bar{\mathbb{Q}}$, with ℓ outside a finite set, such that $\bar{\rho}_{f,\lambda}$ is irreducible. We fix a prime $q \nmid N\ell$ such that the following two conditions hold:

- $q^i \neq 1 \pmod{\ell}$ for $i = 1, \dots, 4$;
- $\bar{\rho}_{f,\lambda}(\mathrm{Frob}_q)$ has a fixed vector.

Here Frob_q is a fixed geometric Frobenius in the Galois group of $\mathbb{Q}$. Then, by Theorem D, there exists a cuspidal automorphic representation Π of PGSp(4) such that Π_∞ is the cohomological holomorphic discrete series representation, Π_p is unramified and tempered for $p \nmid Nq$, and:

- $\rho_{\Pi,\lambda}$ is irreducible, but $\bar{\rho}_{\Pi,\lambda} \simeq \bar{\rho}_{f,\lambda} \oplus \bar{\omega}_\ell^{-1} \oplus \bar{\omega}_\ell^{-2}$;
- Π_q is of type IIIa (hence tempered, generic and ramified);
- Π_p is para-spherical for all primes p dividing N.

Recall that $\rho_{\Pi,\lambda}$ is the four-dimensional λ-adic representation associated to the form Π by Weissauer and Laumon, and that $\bar{\rho}_{\Pi,\lambda}$ is its reduction modulo λ.

In the following, we let V denote the space of $\rho_{\Pi,\lambda}$. This is a four-dimensional vector space over the ℓ-adic field L. We let $\mathcal{O}$ be the ring of integers in L. By an abuse of notation, we let λ also denote the maximal ideal it generates in $\mathcal{O}$. Moreover, $\mathbb{F}$ denotes the residue field of $\mathcal{O}$.

6.2 Choosing a lattice

Under our assumptions, $\bar{\rho}_{\Pi,\lambda}(\mathrm{Frob}_q)$ has eigenvalues $\{1, q, q^2, q^3\}$. By banality, they are all distinct; therefore, by Hensel's lemma, $\rho_{\Pi,\lambda}(\mathrm{Frob}_q)$ has eigenvalues $\{\alpha, \beta, \gamma, \delta\}$ reducing to $\{1, q, q^2, q^3\}$ modulo λ. Let $v \in V$ be an eigenvector for $\alpha \equiv 1 \pmod{\lambda}$. Consider the module it generates,

$$\Lambda = \mathcal{O}[\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})] \cdot v.$$

This is a non-zero Galois-stable cyclic $\mathcal{O}$-module. By the irreducibility of $\rho_{\Pi,\lambda}$, we must have $\Lambda \otimes L = V$, implying that Λ is a Galois-stable $\mathcal{O}$-lattice in V. We look at its reduction $\Lambda_{\mathbb{F}}$; this is cyclic, generated by the class of v. Hence $\bar{\rho}_{f,\lambda}$ is the unique irreducible quotient of $\Lambda_{\mathbb{F}}$.

6.3 Kummer theory

It is known from Kummer theory that $H^1_f(\mathbb{Q}, \omega_\ell) = 0$ (as it should be since $\zeta(0)$ is non-zero). Here, as in [Bel02, Proposition 5.1], we observe that the analogous statement modulo ℓ is true.

Lemma 6.1. $H^1_f(\mathbb{Q}, \bar{\omega}_\ell) = 0$.

Proof. The connecting homomorphism for the Kummer sequence yields

$$H^1(\mathbb{Q}, \bar{\omega}_\ell) \simeq \mathbb{Q}^*/\mathbb{Q}^{*\ell}.$$

Fix $a \in \mathbb{Q}^*$ and let $\delta(a)$ be the corresponding cohomology class. Clearly, $\delta(a)$ is unramified at $p \neq \ell$ if and only if ℓ divides $\mathrm{ord}_p(a)$. By the discussion in Rubin's book [Rub00, p. 26], it is also true that $\delta(a)$ restricts to a class in $H^1_f(\mathbb{Q}_\ell, \bar{\omega}_\ell)$ if and only if ℓ divides $\mathrm{ord}_\ell(a)$. Therefore $H^1_f(\mathbb{Q}, \bar{\omega}_\ell) = 0$. □

6.4 Kato's result

Since $\rho_{f,\lambda}(1)$ has positive weight, the Bloch–Kato conjecture predicts that

$$H^1_f(\mathbb{Q}, \rho_{f,\lambda}(1)) = 0,$$

which was proved in [Kat04]. From this, we deduce the analogous result mod ℓ.

Lemma 6.2. $H^1_f(\mathbb{Q}, \bar{\rho}_{f,\lambda}(1)) = 0$ *for almost all* ℓ.

Proof. In this proof, take V to be the space of $\rho_{f,\lambda}(1)$ and let Λ be a Galois-stable lattice in V. Let W denote the quotient V/Λ. By [Rub00, p. 22, Lemma 1.5.4], there is a natural surjection

$$H^1_f(\mathbb{Q}, \Lambda/\lambda\Lambda) \twoheadrightarrow H^1_f(\mathbb{Q}, W)[\lambda].$$

This is, in fact, an isomorphism for almost all ℓ, since $H^0(\mathbb{Q}, W)=0$ by [Kat04, p. 241, Proposition 14.11]. However, $H^1_f(\mathbb{Q}, W)=0$ by [Kat04, p. 235, Theorem 14.2]. □

6.5 Existence of certain submodules

In this section, we show that $\bar{\omega}_\ell^{-2}$ embeds in $\Lambda_{\mathbb{F}}$. Suppose it does not; then $\bar{\omega}_\ell^{-1}$ would be the unique irreducible subrepresentation of $\Lambda_{\mathbb{F}}$. Writing down a composition series, we get non-split extensions

$$0 \to \bar{\omega}_\ell \to X \to \mathbf{1} \to 0 \quad \text{and} \quad 0 \to \bar{\rho}_{f,\lambda}(1) \to Y \to \mathbf{1} \to 0.$$

Up to a Tate twist, X and $Y^\vee$ are subquotients of $\Lambda_{\mathbb{F}}$. By Lemmas 6.1 and 6.2, to get a contradiction it suffices to show that one of the corresponding cohomology classes lies in the Selmer group.

Lemma 6.3. *X and Y are both Fontaine–Laffaille at ℓ.*

Proof. Since Π is neither CAP nor endoscopic, it follows from [Wei05] that $\rho_{\Pi,\lambda}$ is the representation on the Π_f^K-isotypic component of the étale intersection cohomology (for the middle perversity),

$$IH^3_{\text{et}}(\bar{S}_K \times_{\mathbb{Q}} \bar{\mathbb{Q}}, \bar{\mathbb{Q}}_\ell).$$

Here K is paramodular at primes dividing N, Klingen at q, and hyperspecial outside Nq. Moreover, $\bar{S}_K$ denotes the Satake compactification of the Siegel threefold S_K. The latter has good reduction at $\ell \nmid Nq$, so $\rho_{\Pi,\lambda}$ is crystalline, with Hodge–Tate weights contained in $\{0, 1, 2, 3\}$; see [SU06, p. 41]. Now, X and $Y^\vee$ are both torsion subquotients of $\rho_{\Pi,\lambda}(2)$. If $\ell-1$ is bigger than the Hodge–Tate weights, i.e. if $\ell > 5$, then X and $Y^\vee$ are Fontaine–Laffaille by [BM02, Theorem 3.1.3.3]. □

From the theory of Fontaine and Laffaille [FL82], reviewed by Breuil and Messing in [BM02], it follows that the above extensions are reductions of lattices in certain crystalline representations; see [BM02, Theorems 3.1.3.2 and 3.1.3.3].

Now consider a prime $p \neq \ell$. Clearly, X and Y are then both unramified at $p \nmid Nq$. We are thus left with the two cases $p|N$ and $p = q$. In the first case, we need to show exactness of

$$0 \to \bar{\rho}_{f,\lambda}(1)^{I_p} \to Y^{I_p} \to \mathbf{1} \to 0,$$

where I_p is the inertia group at p, and similarly for X. This requires our minimality assumption that $\bar{\rho}_{f,\lambda}$ is ramified at all primes $p|N$. Moreover, we need to appeal to Conjecture 1.

Lemma 6.4. *Conjecture 1 implies that X and Y have good reduction at $p|N$.*

Proof. Let us first consider X. We need to show that it is unramified at $p|N$. Since $\bar{\rho}_{f,\lambda}$ is the unique irreducible quotient of $\Lambda_{\mathbb{F}}$, the quotient of $\Lambda_{\mathbb{F}}(2)$ by X is $\bar{\rho}_{f,\lambda}(2)$. Therefore, we have inequalities

$$3 - \dim X^{I_p} \leq \dim V^{I_p} - \dim X^{I_p} \leq \dim \Lambda_{\mathbb{F}}^{I_p} - \dim X^{I_p} \leq \dim \bar{\rho}_{f,\lambda}^{I_p} \leq 1.$$

The first inequality follows from Conjecture 1, and the last is our minimality assumption. It follows that X is unramified. Next, let us consider Y. Here the dual of the quotient of $\Lambda_{\mathbb{F}}$ by $\bar{\omega}_\ell^{-1}$ equals $Y(2)$. By the same arguments as before, we then have the following string of inequalities:

$$2 \leq \dim V^{I_p} - 1 \leq \dim \Lambda_{\mathbb{F}}^{I_p} - 1 \leq \dim Y^{I_p} \leq \dim \bar{\rho}_{f,\lambda}^{I_p} + 1 \leq 2.$$

We conclude that all of these inequalities are in fact equalities, so $\dim Y^{I_p} = 2$. □

To get a contradiction, it now suffices to show that X *or* Y is unramified at q. For this, we invoke a result of Genestier and Tilouine [GT05] on the order of the monodromy operator.

Lemma 6.5. *X or Y is unramified at q.*

Proof. In this proof, let N be the monodromy operator on V at q. From Appendix B, we see that Π_q has a unique line fixed by the Klingen parahoric, since it is of type IIIa. Then [GT05, p. 12, Theorem 2.2.5(1)] tells us that $N^2 = 0$. The operator preserves Λ and the composition series of $\Lambda_{\mathbb{F}}$. Suppose X and Y are both ramified. Then, in some basis, N has the form

$$N \sim \begin{pmatrix} 0 & e & a & b \\ & 0 & c & d \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}$$

with e and (c, d) non-zero. However, this cannot happen since $N^2 = 0$. □

This contradicts Lemmas 6.1 and 6.2; therefore $\bar{\omega}_\ell^{-2}$ *does* embed into $\Lambda_{\mathbb{F}}$.

6.6 Proof of Theorem E

Embed $\bar{\omega}_\ell^{-2}$ as a submodule of $\Lambda_{\mathbb{F}}$, and extend it to a composition series. After twisting, this gives an extension

$$0 \to \bar{\rho}_{f,\lambda}(2) \to Z \to \mathbf{1} \to 0,$$

which is non-split since $\bar{\rho}_{f,\lambda}$ is the unique irreducible quotient of $\Lambda_{\mathbb{F}}$. Up to a twist, $Z^\vee$ is the quotient of $\Lambda_{\mathbb{F}}$ by $\bar{\omega}_\ell^{-2}$. The exact same arguments as in the previous section then show that Z has good reduction away from q (assuming Conjecture 1). It remains to deal with the prime q.

Lemma 6.6. *Z is unramified at q.*

Proof. The extension Z determines a cohomology class in $H^1(\mathbb{Q}_q, \bar{\rho}_{f,\lambda}(2))$. Let c be a cocycle representing this class. Since $\bar{\rho}_{f,\lambda}(2)$ is unramified at q, the cocycle restricts to a homomorphism from the inertia group I_q to the space of $\bar{\rho}_{f,\lambda}(2)$. As $q \neq \ell$, it obviously factors through the tame quotient; indeed, it factors through the homomorphism $t_\ell : I_q \twoheadrightarrow \mathbb{Z}_\ell$. Recall from [Tat79, p. 21] that

$$t_\ell(\mathrm{Frob}_q^{-1} \cdot \sigma \cdot \mathrm{Frob}_q) = q \cdot t_\ell(\sigma)$$

for $\sigma \in I_q$. Clearly, the left-hand side is independent of the choice of a Frobenius Frob_q in the Galois group of $\mathbb{Q}_q$. We then immediately deduce an analogous relation satisfied by c. Now we invoke the cocycle relation satisfied by c. Using it twice, we find that

$$c(\mathrm{Frob}_q^{-1} \cdot \sigma \cdot \mathrm{Frob}_q) = c(\mathrm{Frob}_q^{-1}) + \mathrm{Frob}_q^{-1} \cdot c(\sigma \cdot \mathrm{Frob}_q) = \mathrm{Frob}_q^{-1} \cdot c(\sigma)$$

for $\sigma \in I_q$, since $\bar{\rho}_{f,\lambda}$ is unramified at q, and

$$c(\mathrm{Frob}_q^{-1}) = -\mathrm{Frob}_q^{-1} \cdot c(\mathrm{Frob}_q).$$

The action of Frob_q^{-1} on the vector $c(\sigma)$ is given by the Tate twist $\bar{\rho}_{f,\lambda}(2)$:

$$\mathrm{Frob}_q^{-1} \cdot c(\sigma) = \bar{\rho}_{f,\lambda}(\mathrm{Frob}_q^{-1}) \cdot q^2 c(\sigma).$$

Consequently, we end up with the identity

$$\bar{\rho}_{f,\lambda}(\mathrm{Frob}_q) \cdot c(\sigma) = q \cdot c(\sigma).$$

Therefore, if $c(\sigma) \neq 0$ for some $\sigma \in I_q$, we see that $c(\sigma)$ is an eigenvector for $\bar{\rho}_{f,\lambda}(\mathrm{Frob}_q)$ with eigenvalue q. However, the eigenvalues of $\bar{\rho}_{f,\lambda}(\mathrm{Frob}_q)$ are $\{1, q^3\}$ by assumption. □

This finishes the proof of Theorem E.

Acknowledgements

This paper is based on my doctoral dissertation, completed at the California Institute of Technology. I would to express sincere gratitude and appreciation to my advisor D. Ramakrishnan, both for suggesting the problem about level-raising congruences for inner forms of GSp(4) and for sharing his insights at many stages of the project. In addition, many thanks are due to C. Skinner and A. Wiles for their encouragement. Finally, I would like to acknowledge the impact of the ideas of J. Bellaiche, L. Clozel, K. Ribet, and R. Taylor.

Appendix A. Iwahori-spherical representations

Table A1 is essentially [Sch05, Table 1]; it is included it here for the reader's convenience, and we are grateful to Ralf Schmidt for his permission to do so. Throughout, we use the notation

Table A1. Iwahori-spherical representations of GSp(4).

		Constituent of	Representation	Tempered	L^2	Generic
I		$\chi_1 \times \chi_2 \rtimes \sigma$	$\chi_1 \times \chi_2 \rtimes \sigma$	$\lvert\chi_i\rvert = \lvert\sigma\rvert = 1$		•
II	a	$\nu^{1/2}\chi \times \nu^{-1/2}\chi \rtimes \sigma$,	$\chi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$	$\lvert\chi\rvert = \lvert\sigma\rvert = 1$		•
	b	$\chi^2 \notin \{\nu^{\pm 1}, \nu^{\pm 3}\}$	$\chi\mathbf{1}_{\mathrm{GL}(2)} \rtimes \sigma$			
III	a	$\chi \times \nu \rtimes \nu^{-1/2}\sigma$,	$\chi \rtimes \sigma\mathrm{St}_{\mathrm{GL}(2)}$	$\lvert\chi\rvert = \lvert\sigma\rvert = 1$		•
	b	$\chi \notin \{\mathbf{1}, \nu^{\pm 2}\}$	$\chi \rtimes \sigma\mathbf{1}_{\mathrm{GL}(2)}$			
IV	a	$\nu^2 \times \nu \rtimes \nu^{-3/2}\sigma$	$\sigma\mathrm{St}_{\mathrm{GSp}(4)}$	•	•	•
	b		$L((\nu^2, \nu^{-1}\sigma\mathrm{St}_{\mathrm{GL}(2)}))$			
	c		$L((\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-3/2}\sigma))$			
	d		$\sigma\mathbf{1}_{\mathrm{GSp}(4)}$			
V	a	$\nu\xi_0 \times \xi_0 \rtimes \nu^{-1/2}\sigma$,	$\delta([\xi_0, \nu\xi_0], \nu^{-1/2}\sigma)$	•	•	•
	b	$\xi_0^2 = \mathbf{1}$, $\xi_0 \neq \mathbf{1}$	$L((\nu^{1/2}\xi_0\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma))$			
	c		$L((\nu^{1/2}\xi_0\mathrm{St}_{\mathrm{GL}(2)}, \xi_0\nu^{-1/2}\sigma))$			
	d		$L((\nu\xi_0, \xi_0 \rtimes \nu^{-1/2}\sigma))$			
VI	a	$\nu \times \mathbf{1} \rtimes \nu^{-1/2}\sigma$	$\tau(S, \nu^{-1/2}\sigma)$	•		•
	b		$\tau(T, \nu^{-1/2}\sigma)$	•		
	c		$L((\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma))$			
	d		$L((\nu, \mathbf{1} \rtimes \nu^{-1/2}\sigma))$			

Table B1. Dimensions of the parahoric fixed spaces.

		Representation	Remarks	K	$\tilde{K}$	J_α	J_β	I
I		$\chi_1 \times \chi_2 \rtimes \sigma$		1	2	4	4	8
II	a	$\chi \mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$		0	1	2	1	4
	b	$\chi \mathbf{1}_{\mathrm{GL}(2)} \rtimes \sigma$		1	1	2	3	4
III	a	$\chi \rtimes \sigma \mathrm{St}_{\mathrm{GL}(2)}$		0	0	1	2	4
	b	$\chi \rtimes \sigma \mathbf{1}_{\mathrm{GL}(2)}$		1	2	3	2	4
IV	a	$\sigma \mathrm{St}_{\mathrm{GSp}(4)}$		0	0	0	0	1
	b	$L((\nu^2, \nu^{-1}\sigma \mathrm{St}_{\mathrm{GL}(2)}))$	not unitary	0	0	1	2	3
	c	$L((\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-3/2}\sigma))$	not unitary	0	1	2	1	3
	d	$\sigma \mathbf{1}_{\mathrm{GSp}(4)}$	irrelevant	1	1	1	1	1
V	a	$\delta([\xi_0, \nu\xi_0], \nu^{-1/2}\sigma)$		0	0	1	0	2
	b	$L((\nu^{1/2}\xi_0 \mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma))$		0	1	1	1	2
	c	$L((\nu^{1/2}\xi_0 \mathrm{St}_{\mathrm{GL}(2)}, \xi_0\nu^{-1/2}\sigma))$		0	1	1	1	2
	d	$L((\nu\xi_0, \xi_0 \rtimes \nu^{-1/2}\sigma))$		1	0	1	2	2
VI	a	$\tau(S, \nu^{-1/2}\sigma)$		0	0	1	1	3
	b	$\tau(T, \nu^{-1/2}\sigma)$		0	0	0	1	1
	c	$L((\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma))$		0	1	1	0	1
	d	$L((\nu, \mathbf{1} \rtimes \nu^{-1/2}\sigma))$		1	1	2	2	3

of [ST93]. Let B be the Borel subgroup of upper triangular matrices in GSp(4). Let P_α and P_β be the maximal parabolic subgroups containing B. Their matrix realizations are given in § 2.3. If χ_1, χ_2 and σ are characters of GL(1), we denote by $\chi_1 \times \chi_2 \rtimes \sigma$ the representation of GSp(4) obtained by normalized induction from the following character of B:

$$\begin{pmatrix} a & & & \\ & b & & \\ & & cb^{-1} & \\ & & & ca^{-1} \end{pmatrix} \mapsto \chi_1(a)\chi_2(b)\sigma(c).$$

Similarly, if τ is a representation of GL(2), we let $\tau \rtimes \sigma$ be the representation of GSp(4) obtained by normalized induction from

$$\begin{pmatrix} g & \\ & c\,{}^\tau g^{-1} \end{pmatrix} \mapsto \sigma(c)\tau(g).$$

Moreover, $\sigma \rtimes \tau$ denotes the representation induced from the Klingen parabolic,

$$\begin{pmatrix} c & & \\ & g & \\ & & c^{-1}\det g \end{pmatrix} \mapsto \sigma(c)\tau(g).$$

In Table A1, $\nu = |\cdot|$ is the normalized absolute value, χ_0 is the unique non-trivial unramified quadratic character, St is the Steinberg representation, $\mathbf{1}$ is the trivial representation, and $L((-))$ denotes the unique irreducible quotient when it exists.

Appendix B. Parahoric fixed spaces

Table B1 is essentially [Sch05, Table 3]. Here K is hyperspecial, $\tilde{K}$ is paramodular, I is Iwahori, and J_α and J_β denote the Klingen- and Siegel-parahoric subgroups, respectively. For example, J_α is the inverse image of P_α over the residue field under the natural reduction map.

References

Art96 J. Arthur, *On local character relations*, Selecta Math. **2** (1996), 501–579.

Art98 J. Arthur, *Towards a stable trace formula*, in *Proceedings of the international congress of mathematicians (Berlin, 1998), vol. II*, Doc. Math., extra vol. ICM II(1998), 507–517.

Bel02 J. Bellaiche, *Congruences endoscopiques et représentations Galoisiennes.* PhD thesis, Université Paris XI, Orsay (2002) (http://people.brandeis.edu/∼jbellaic/).

BR94 D. Blasius and J. Rogawski, *Zeta functions of Shimura varieties*, in *Motives (Seattle, WA, 1991)*, Proceedings of Symposia in Pure Mathematics, vol. 55, part 2 (American Mathematical Society, Providence, RI, 1994), 525–571.

BK90 S. Bloch and K. Kato, *L-functions and Tamagawa numbers of motives*, in *The Grothendieck Festschrift, vol. I*, Progress in Mathematics, vol. 86 (Birkhäuser, Boston, MA, 1990), 333–400.

BM02 C. Breuil and W. Messing, *Torsion étale and crystalline cohomologies*, in *Cohomologies p-adiques et applications arithmétiques, II*, Astérisque, vol. 279 (Société Mathématique de France, Paris, 2002), 81–124.

Car79 P. Cartier, *Representations of p-adic groups: a survey*, in *Automorphic forms, representations and L-functions*, Proceedings of Symposia in Pure Mathematics, vol. 33, part 1 (American Mathematical Society, Providence, RI, 1979), 111–155.

Cas81 W. Casselman, *A new nonunitarity argument for p-adic representations*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **28** (1982), 907–928.

CN90 C.-L. Chai and P. Norman, *Bad reduction of the Siegel moduli scheme of genus two with* $\Gamma_0(p)$-*level structure*, Amer. J. Math. **112** (1990), 1003–1071.

Clo00 L. Clozel, *On Ribet's level-raising theorem for* $U(3)$, Amer. J. Math. **122** (2000), 1265–1287.

deJ93 J. de Jong, *The moduli spaces of principally polarized abelian varieties with* $\Gamma_0(p)$-*level structure*, J. Algebraic Geom. **2** (1993), 667–688.

Del71 P. Deligne, *Formes modulaires et représentations* ℓ-*adiques*, in *Séminaire Bourbaki 1968/69*, Lecture Notes in Mathematics, vol. 79 (Springer, Berlin, 1971), 347–363.

DS74 P. Deligne and J.-P. Serre, *Formes modulaires de poids* 1, Ann. Sci. Ecole Norm. Sup. (4) **7** (1974), 507–530.

FL82 J.-M. Fontaine and G. Laffaille, *Construction de représentations p-adiques*, Ann. Sci. Ecole Norm. Sup. (4) **15** (1982), 547–608.

Gan08 W. T. Gan, *The Saito-Kurokawa space of* PGSp_4 *and its transfer to inner forms*, in *Proceedings of the AIM workshop "Eisenstein Series and Applications" (Palo Alto, 2005)*, Progress in Mathematics, vol. 258 (Birkhäuser, Boston, MA, 2008), (http://www.math.ucsd.edu/∼wgan/.)

GG05 W. T. Gan and N. Gurevich, *Non-tempered Arthur packets of* G_2, in *Automorphic representations, L-functions and applications: progress and prospects*, Ohio State Univ. Math. Res. Inst. Publ., vol. 11 (de Gruyter, Berlin, 2005), 129–155.

GT05 A. Genestier and J. Tilouine, *Systemes de Taylor-Wiles pour* GSp_4, in *Formes automorphes. II: Le cas du groupe* GSp(4), Astérisque, no. 302 (Société Mathématique de France, Paris, 2005), 177–290.

Gre63 M. Greenberg, *Schemata over local rings. II*, Ann. of Math. (2) **78** (1963), 256–266.

HN02 T. Haines and B. C. Ngô, *Nearby cycles for local models of some Shimura varieties*, Compositio Math. **133** (2002), 117–150.

Hal89 T. Hales, *Shalika germs on* GSp(4), in *Orbites unipotentes et representations. II*, Astérisque, no. 171–172 (Société Mathématique de France, Paris, 1989), 195–256.

Hal97 T. Hales, *The fundamental lemma for* Sp(4), Proc. Amer. Math. Soc. **125** (1997), 301–308.

JL89 B. Jordan and R. Livne, *Conjecture "epsilon" for weight* $k > 2$, Bull. Amer. Math. Soc. **21** (1989), 51–56.

Kat04 K. Kato, *p-adic Hodge theory and values of zeta functions of modular forms*, in *Cohomologies p-adiques et applications arithmetiques. III*, Asterisque, no. 295 (Société Mathématique de France, Paris, 2004), 117–290.

Kaz87 D. Kazhdan and G. Lusztig, *Proof of the Deligne–Langlands conjecture for Hecke algebras*, Invent. Math. **87** (1987), 153–215.

Kot83 R. Kottwitz, *Sign changes in harmonic analysis on reductive groups*, Trans. Amer. Math. Soc. **278** (1983), 289–297.

Kot86 R. Kottwitz, *Base change for unit elements of Hecke algebras*, Compositio Math. **60** (1986), 237–250.

Kud86 S. Kudla, *On the local theta-correspondence*, Invent. Math. **83** (1986), 229–255.

Kud S. Kudla, *Notes on the local theta-correspondence*, unpublished notes (http://www.math.utoronto.ca/∼skudla/ssk.research.html).

KRS92 S. Kudla, S. Rallis and D. Soudry, *On the degree* 5 *L-function for* Sp(2), Invent. Math. **107** (1992), 483–541.

LM04 J.-P. Labesse and W. Muller, *Weak Weyl's law for congruence subgroups*, Asian J. Math. **8** (2004), 733–745.

Lan76 R. Langlands, *On the functional equations satisfied by Eisenstein series*, Lecture Notes in Mathematics, vol. 544 (Springer, Berlin, New York, 1976).

LS90 R. Langlands and D. Shelstad, *Descent for transfer factors. The Grothendieck Festschrift, vol. II*, Progress in Mathematics, vol. 87 (Birkhäuser, Boston, MA, 1990), 485–563.

Lau97 G. Laumon, *Sur la cohomologie a supports compacts des varietes de Shimura pour* $\mathrm{GSp}(4)_{\mathbb{Q}}$, Compositio Math. **105** (1997), 267–359.

Laz00 X. Lazarus, *Module universel non ramifie pour un groupe reductif p-adique.* PhD thesis, Université Paris XI, Orsay (2000).

Li90 J.-S. Li, *Theta lifting for unitary representations with nonzero cohomology*, Duke Math. J. **61** (1990), 913–937.

Pia83 I. Piatetski-Shapiro, *On the Saito–Kurokawa lifting*, Invent. Math. **71** (1983), 309–338.

Ram08 D. Ramakrishnan, *Irreducibility and cuspidality*, in *Representation theory and automorphic forms*, Progress in Mathematics, vol. 255 (Birkhäuser, Boston, MA, 2008), 1–27.

Rap90 M. Rapoport, *On the bad reduction of Shimura varieties*, in *Automorphic forms, Shimura varieties, and L-functions (Ann Arbor, MI, 1988), vol. II*, Perspectives in Mathematics, vol. 11 (Academic Press, Boston, MA, 1990), 253–321.

Rib85 K. Ribet, *On ℓ-adic representations attached to modular forms. II*, Glasgow Math. J. **27** (1985), 185–194.

Rob96 B. Roberts, *The theta correspondence for similitudes*, Israel J. Math. **94** (1996), 285–317.

Rob99 B. Roberts, *The non-Archimedean theta correspondence for* GSp(2) *and* GO(4), Trans. Amer. Math. Soc. **351** (1999), 781–811.

Rub00 K. Rubin, *Euler systems (Hermann Weyl lectures, The Institute for Advanced Study)*, Annals of Mathematics Studies, vol. 147 (Princeton University Press, Princeton, NJ, 2000).

ST93 P. Sally and M. Tadic, *Induced representations and classifications for* $\mathrm{GSp}(2, F)$ *and* $\mathrm{Sp}(2, F)$, Mem. Soc. Math. France **52** (1993), 75–133.

Sch05 R. Schmidt, *Iwahori-spherical representations of* GSp(4) *and Siegel modular forms of degree 2 with square-free level*, J. Math. Soc. Japan **57** (2005), 259–293.

She79 D. Shelstad, *Characters and inner forms of a quasi-split group over* $\mathbb{R}$, Compositio Math. **39** (1979), 11–45.

She82 D. Shelstad, *L-indistinguishability for real groups*, Math. Ann. **259** (1982), 385–430.

SU06 C. Skinner and E. Urban, *Sur les déformations p-adiques de certaines représentations automorphes*, J. Inst. Math. Jussieu **5** (2006), 629–698.

Sor06 C. Sorensen, *A generalization of level-raising congruences for algebraic modular forms*, Ann. Inst. Fourier (Grenoble) **56** (2006), 1735–1766.

Sou88 D. Soudry, *The CAP representations of* GSp(4, $\mathbb{A}$), J. Reine Angew. Math. **383** (1988), 87–108.

Tat79 J. Tate, *Number theoretic background*, in *Automorphic forms, representations and L-functions (Corvallis, OR, 1977)*, Proceedings of Symposia in Pure Mathematics, vol. 33, part 2 (American Mathematical Society, Providence, RI, 1979), 3–26.

Tay89 R. Taylor, *On Galois representations associated to Hilbert modular forms*, Invent. Math. **98** (1989), 265–280.

Tit79 J. Tits, *Reductive groups over local fields*, in *Automorphic forms, representations and L-functions*, Proceedings of Symposia in Pure Mathematics, vol. 33, part 1 (American Mathematical Society, Providence, RI, 1979), 29–69.

Vig96 M.-F. Vigneras, *Representations ℓ-modulaires d'un groupe reductif p-adique avec $\ell \neq p$*, Progress in Mathematics, vol. 137 (Birkhäuser, Boston, MA, 1996).

Wal97 J.-L. Waldspurger, *Le lemme fondamental implique le transfert*, Compositio Math. **105** (1997), 153–236.

Wei05 R. Weissauer, *Four dimensional Galois representations*, in *Formes automorphes. II: Le cas du groupe* GSp(4), Astérisque, no. 302 (Société Mathématique de France, Paris, 2005), 67–150.

Wei1 R. Weissauer, *Character identities and Galois representations related to the group* GSp(4), Preprint (http://www.mathi.uni-heidelberg.de/∼weissaue/papers.html).

Wei2 R. Weissauer, *Endoscopy for* GSp(4), Preprint (http://www.mathi.uni-heidelberg.de/∼weissaue/papers.html).

Claus M. Sorensen claus@princeton.edu
Department of Mathematics, Princeton University, Princeton, NJ 08544, USA